The Development and Sensitivity Analysis of the 2010 Demographic Analysis Estimates

by

The Demographic Analysis Research Team
U.S. Census Bureau
12/16/10

This paper reflects the work of the U.S. Census Bureau's 2010 Demographic Analysis Research Team. Members of the Demographic Analysis Research Team include: J. Gregory Robinson, Victoria Velkoff, Jason Devine, Kirsten West, Renuka Bhaskar, Melissa Scopilliti, Bethany DeSalvo, Jennifer Ortman, David Armstrong, Antonio Bruce, Tammany Burden, Katherine Condon, Rachel Cortes, Christopher Dick, David Dixon, Belkines arenas Germosen, Sarah Gibb, Elizabeth Grieco, Christine Guarneri, Gregory Harper, Frederick Hollmann, Eric Jensen, Peter Johnson, Rodger Johnson, Alexa Kennedy Jones-Puthoff, Luke Larsen, Yeris Mayol-Garcia, Caleb Miller, Larry Sink, Tiffany Thompson, Amin Vafa, Grayson Vincent, and David Waddington. The team would like to thank the following people for their contributions: Howard Hogan, Enrique Lamas, Frank Hobbs, Barbara Anderson, Irma Elo, Jennifer Van Hook, Joseph Salvo, and Jeffrey Passel.

INTRODUCTION

In the United States, a collection of methods have historically been used to develop estimates of the population for comparison with decennial census counts. These estimates are developed from various types of demographic data in order to build a historical accounting of population change. The term Demographic Analysis (DA) has been used to refer to this approach for assessing the quality of the census. For the 2010 Census, the U.S. Census Bureau is again using the DA estimates along with results from operational indicators and a post-enumeration survey to assess the quality of the decennial census.[1]

In past decades, the Census Bureau has produced the DA estimates at the national level for single-year birth cohorts by sex and two broad race categories, Black and non-Black. Reflecting the contributions of many demographers, two distinct methodologies have evolved over time for separately estimating the population under 65 years of age and the population 65 years and over. Administrative records on births, deaths, and estimates of international migration are used to estimate the population under 65. Estimates of the population 65 and over are developed from data on Medicare enrollment and estimates of the number of those not enrolled. The DA estimates have been limited to the national level and the Black and non-Black race categories because of the limitations of the available historical data.[2, 3]

This paper describes the development of the 2010 DA population estimates. In order to show the sensitivity of the DA estimates to the use of different but plausible input values, alternative estimates for each component of the DA methodology were developed. Specific combinations of alternative inputs were then used to produce five series of DA estimates by single year of age, sex,

[1] In addition to various operational indicators, the Census Bureau has historically relied on two principal methods to provide measures of the quality of each census. One method is Demographic Analysis, which is the topic of this paper. The other method is based on a post-enumeration survey and dual-system estimation.

[2] For estimates to be produced for areas below the national level using the DA methodology a source for measuring internal migration by race starting in 1945 would be needed. A data source suitable for this purpose does not exist.

[3] Because of the limited race detail available from historical vital statistics data, the DA estimates have traditionally been developed by single year of age and sex for only two race categories, Black and non-Black. Black is used throughout this report to refer to the Black or African American population. Because of the uncertainty that "classification" error would introduce into estimates for race categories other than Black and White, all races other than Black are combined to form a non-Black category.

and the DA race categories for use in the demographic analysis of the 2010 Census counts.[4] The specific combinations of alternative estimates were selected to communicate the uncertainty around the DA estimates and include a plausible low series, three middle series, and a plausible high series. These estimates will provide an essentially independent basis for assessing the accuracy of the 2010 Census counts.

Overview of the DA Methodology

DA represents a macro-level approach for analyzing census counts that relies on comparisons with national-level estimates to provide information on the quality of the census. Historically, it was assumed that differences represented error in the census: a lower census count was assumed to reflect a net undercount, and a higher census count indicated a net overcount. The DA estimates were used to provide a measure of net coverage error and did not provide information on the separate effects of the different types of coverage error (omissions, erroneous inclusions, or duplicates) or content error. For DA in 2010, differences between the DA estimate and the census count are not assumed to represent error in the census, but are referred to as differences.

Separate methods are used in DA to obtain estimates for two segments of the total population: the population under age 65, and the population aged 65 and over:

(1) Ages Under 65.

The DA estimates for the population aged 0 to 64 (P_{0-64}) are derived by the basic demographic accounting equation applied to each birth cohort:

$$P_{0-64} = B - D + I - E \qquad (1)$$

DA estimates for the population below age 65 are developed from a compilation of historical estimates of the components of population change (the cohort-component approach): births corrected for underregistration, beginning with April 1, 1945 (B); deaths to persons born since April 1, 1945 (D); immigrants below age 65 (I); and emigrants below age 65 (E).

[4] Each set of DA estimates by single year of age, sex, and the DA race and ethnicity categories is referred to here as a "series" of estimates.

For DA in 2000, the cohort-component approach was not used for populations born prior to 1935 because the birth registration system did not include all states until 1933. For DA in 2010, while the cohort-component approach is used as the primary estimate of the population below age 65, data from 1935 to 1944 were carried forward to produce an estimate of the population aged 65 to 74. This results in an overlap between the estimates developed using the component approach and the Medicare-based estimates. The differences between these two series for ages 65 to 74 have been examined, and the results are discussed later in this paper.

In 2000, births (234.9 million) represented by far the largest component in equation 1. The immigration component (32.6 million) was second largest, followed by deaths (14.8 million) and emigrants (5.5 million) (Robinson, 2010). The number of deaths is small relative to births because of the lower rate of mortality for the population under 65. For DA in 2010, births (249.9 million) continue to be the largest component of the DA estimates. As in 2000, the immigration component is the second largest, followed by deaths and emigration.[5] The DA estimates produced in 2000 represent the starting point for the development of the 2010 DA estimates for the population under age 65. To obtain the DA estimates for April 1, 2010, births and immigration were added, and deaths and emigration were subtracted. Key assumptions for the use of the vital statistics data and the estimates of international migration from DA in 2000 were also either revisited or altered for the DA 2010 estimates.

The actual calculations used to develop the DA estimates are carried out for single-year birth cohorts by sex and race (Black and non-Black), and ethnicity (Hispanic and non-Hispanic). For example, the estimate of the population aged 58 on April 1, 2010 was based on births from April of 1951 through March of 1952 (corrected for underregistration), reduced by deaths to the cohort in each year between 1951 and 2010, with additions and subtractions for estimated immigration and emigration of the cohort between 1951 and 2010.

(2) Ages 65 and Over.

[5] The value for births for 2010 was taken from the DA Middle Series estimate.

Administrative data on aggregate Medicare enrollments are used to develop an estimate of the population aged 65 and over (P_{65+}):

$$P_{65+} = M + m \qquad (2)$$

where M is the aggregate Medicare enrollment and m is the estimate of the number not enrolled in Medicare.[6] Although Medicare enrollment is generally presumed to be quite complete, underenrollment factors are applied to account for individuals who are not enrolled. Some groups are not eligible to enroll, such as federal employees who are covered under a specific retirement program; some may delay enrollment until a date later than when they became eligible; and some may never enroll. In 2000, an allowance was made for an estimated 1.3 million not enrolled, or 3.8 percent of the estimated population aged 65 and over. For 2010, an estimated 1.5 million people were not enrolled, or 3.9 percent of the estimated population aged 65 and over. Underenrollment factors are based on estimates of Medicare coverage developed from the Current Population Survey (CPS) and data on age at enrollment in the Medicare file.

The estimates for the population under 65 developed using the cohort-component approach are combined with the Medicare-based estimates for the population 65 and over to produce the total DA population estimate.

Demographic Analysis in 2010

The DA program for 2010 continued the practice of producing estimates by age, sex, and Black and non-Black using the methodology just described, but also included the production of estimates by Hispanic origin for ages under 20. The estimates by Hispanic origin were limited to ages under 20 due to the limitation of the historical estimates of international migration and vital statistics data. These limitations will be discussed in more detail later in this paper. There was also an effort that focused on assessing and clearly communicating the uncertainty in the DA estimates through the production of five series of estimates that incorporated alternative assumptions for each component. In addition to the five series of estimates, additional estimates were produced for each component to further assess the degree of uncertainty around the DA estimates.

[6] Medicare is a federal health insurance program that covers most people aged 65 and over.

While the quality of the vital statistics that serve as the primary component of the DA estimates has improved to virtually complete coverage, there remains uncertainty in the historical birth data. Increases in international migration and changes in the racial makeup of the population have also presented new challenges to the DA methodology over the last two decades. The Census Bureau has responded to the increase in international migration by developing new methods for estimating this component using the latest data sources, including the American Community Survey (ACS).

The increase in the diversity of the U.S. population—specifically, the increase in multiracial births—and changes in how information on race is collected in the census posed a substantial challenge for the DA estimates by Black and non-Black both in 2000 and in 2010. Census 2000 was the first census to include the option to mark more than one race based on the 1997 U.S. Office of Management and Budget (OMB) revisions to the federal standards on collecting information on race and ethnicity. The 2010 Census also includes this option. Jones and Smith (2003) researched the use of Census 2000 responses for recognizing patterns of how interracial/interethnic parents identify their children with regard to race and ethnicity. The approach used for DA in 2010 to classify births as Black or non-Black and Hispanic or non-Hispanic based on the reported race and Hispanic origin of the parents on the birth certificate represents a continuation of this work.

SENSITIVITY ANALYSIS OF THE DA ESTIMATES

The lack of an explicit probability model for the assessment of the accuracy of the DA estimates has been consistently identified as a shortcoming of the DA methodology by statisticians. Demographers have struggled with the development of these measures mainly because the uncertainty in DA derives from judgment about the choice of data sets rather than statistical variation within a given set. The errors in these data are generally not subject to sampling error; therefore, a statistically-based confidence interval is difficult to develop using conventional statistical techniques. When formal uncertainty models have been developed, they have done little to help place the uncertainty around the DA estimates into a meaningful context. For DA in 2010,

the reliance on survey-based estimates of international migration means that a major component of the DA methodology will now include uncertainty due to being based on a sample. This uncertainty can be quantified through the use of probability theory. While a statistician might rely on mathematical models and probability theory to quantify uncertainty, the demographer would tend to rely on implicit models, and treat uncertainty informally, if at all (Hogan et al., 2003). For 2010, a systematic analysis of the DA estimates was conducted that incorporated both statistical measures of uncertainty where applicable (e.g., estimates of international migration), and results from alternative assumptions about each component. Once incorporated, these assumptions were used to develop five series of estimates to more clearly communicate the uncertainty around each DA estimate by single year of age, sex, and the DA race categories.

Uncertainty in each component of DA contributes to the overall uncertainty in the DA estimates in different ways. For example, while the level of uncertainty for each year of births was limited to a specific birth cohort, uncertainty around the estimates of international migration had a greater impact on the uncertainty around the DA estimates for the ages in which international migration is concentrated (e.g., ages 20-39).

In 2000, efforts to assess the uncertainty in the DA estimates focused mainly on reassessing the DA estimates of international migration and the uncertainty around this component. This shift was due to a large initial difference between the DA estimate and the Census 2000 count and may also have been partly due to the acceptance of recent vital statistics data as complete. However, because of the large size of the birth component relative to the other components, uncertainty around the historical birth records also continues to contribute to the overall uncertainty in the DA estimates. Inaccuracies in the estimated completeness of birth registration and the assumption that birth registration was complete starting in 1985 can lead to errors in the estimate of the size of the native-born population.[7] For DA in 2010, the work from the previous efforts was combined with more recent work to provide a more complete assessment of the uncertainty in the DA estimates. While the range of estimates provided as part of the 2010 DA program does not result in a measure of the accuracy of the DA estimate based on an explicit probability model, it should allow for a

[7] The native-born population includes those born in the United States and those born abroad to U.S. citizen parents.

meaningful assessment of the possible range of the DA estimates due to both sampling and nonsampling error.

Development of a range of DA estimates

The main data sources for understanding the size of the U.S. population on April 1, 2000 include Census 2000 counts, estimates of coverage from the 2000 Accuracy and Coverage Evaluation (A.C.E.) Revision II, the 2000 DA estimates, and counts from previous censuses.[8] This information combined with basic assumptions about natural increase (births minus deaths) and international migration can be used to start to establish a plausible range of estimates of the size of the population in 2010. The generally-accepted quality of recent vital-statistics data makes births and deaths a well documented component of change in the size of the U.S. population. Gains from natural increase alone would result in a net increase of approximately 17 million people since 2000. This, combined with a low but plausible estimate of net international migration would provide an estimate that could serve as a plausible low estimate of the April 1, 2010 population. Similarly, the use of what would generally be considered a high but plausible estimate of international migration combined with natural increase could serve as a plausible high estimate of the 2010 population.

The five series of DA estimates developed for April 1, 2010 described in this paper represent a more rigorous version of the simple exercise just described. Various approaches could be used for each component of the DA estimates, each resulting in a different estimate. However, many possible variants of the DA methodology have little impact on the DA estimates. Census Bureau analysts, in consultation with external experts, determined which variants had a large enough impact to be included in the DA estimates.

Expert input was obtained through a conference held on January 8, 2010. This paper was prepared for a second conference scheduled for December 6, 2010 during which five series of DA estimates by single year of age, sex, and the DA race and ethnicity categories will be disseminated. The

[8] A.C.E. refers to the Census 2000 post-enumeration surveys. Post-enumeration surveys rely on case-by-case matching of persons in an independent survey and the dual system estimation methodology to estimate coverage error. The survey-based coverage measurement program associated with the 1980 Census was called the Post-Enumeration Program (PEP); in the 1990 Census it was called the Post-Enumeration Survey (PES).

range of estimates is intended to inform the public of the uncertainty around the DA estimates. The five series of DA estimates will be added to this paper as an addendum for distribution during the December 6 conference. Table 1 provides historical DA estimates and Census counts from 1940 to 2000 by Black and non-Black. The following sections provide an overview of each component of the DA estimates, the major factors contributing to the uncertainty in each component, and the variants used to develop the DA estimates. Table 2 is a matrix of the five DA estimates and the components that were used for each of the five series.

Uncertainty in the vital statistics component

The birth and death data used in DA come from a combination of historical tabulations and micro-level files. The micro-level birth file includes basic information for each birth that occurred from 1980 to 2007. The greatest uncertainty in the vital statistics occurs in the birth records from earlier years which must be combined with estimates of birth underregistration to obtain a complete accounting of births. The use of a separate methodology based on Medicare enrollment for estimating the population 65 and over makes it possible to rely only on births that have occurred from 1945 to 2010. While we must rely on death records for infants and children from all years since 1945, the largest portion of deaths to those under 65 will have occurred in more recent years, and is therefore expected to be of higher quality.

Births

Several issues may lead to errors in the number of registered births. Births might not be reported for a variety of reasons: registration may be delayed and not included in what is provided to the National Center for Health Statistics (NCHS); likewise, processing or administrative errors may lead to the incorrect number being reported to NCHS. Births may also be included by NCHS for a child that should not be included in the census universe.[9]

Three national birth registration tests (BRT) provide the basis for adjusting the reported births for under-registration for use in DA and the assessment of the uncertainty around the registered births.

[9] From 1980 to 2006, just under 131 thousand births occurred to mothers who recorded on the birth certificate that they were not residents of the United States. While these children would be considered native born, in some cases they may not be considered part of the resident population for the 2010 Census. While estimates could have been produced both with and without these births, because of the relatively minor impact on the overall DA estimate, each of the alternative birth series excludes the births to non-residents.

The births used in DA for each year prior to April 1, 1985 are the result of applying registration completeness "correction" factors to each census year of births. Corrections are largest in the years before 1950. For DA in 2000, a total of 3.3 million births were added for the period from 1935 through 1984 based on the correction factors, increasing the total estimated births during this period from 172.4 million to 175.7 million. The 2010 DA estimates only include births starting in April of 1945, which reduces the estimated number of births added due to underregistration from the 3.3 million added to the 2000 DA estimates to 1.5 million.

Based on a review of the DA components during the evaluation of the 2000 DA estimates, birth registration is now assumed to be complete starting in 1985 (which coincides with the first year that birth statistics were reported electronically from all states), and to have remained at 100 percent through 2000 (McDevitt, O'Connell, and Joyce, 2001). For 2010, it was again assumed that birth registration became complete in 1985.

The BRTs were conducted for 1940, 1950, and 1964-1968 (see references). The 1940 and 1950 tests involved matching birth records from children born during a 3-month (1940) and 4-month (1950) period preceding the census with special infant registration cards filled out during the census. The 1964-1968 test included 14,632 cases from the 1969-1970 CPS and Health Interview Survey. The birth registration correction factors for non-test years were derived through interpolation between test dates by in and out of hospital births. A linear interpolation was also used for the years between 1966 and 1985 (first year complete coverage was assumed). The estimated level of completeness was 92.5 percent in 1940 (81.9 percent for Black births) and improved to 97.9 percent in 1950 (93.7 for Black births). By 1964-68, it was estimated that over 99 percent of all births were registered. The proportion of births occurring in and out of hospitals was a large factor in the overall completeness of registration by race and across time. Even in the 1940 test, which showed a large overall difference in registration completeness for the White (94.0 percent) and Black and other races (82.0 percent) categories, there was a much smaller difference between these categories when the birth occurred in a hospital: 98.6 percent compared to 96.3 percent. Out-of-hospital births are estimated to have declined from 44 percent in 1940 to 1 percent in 1969, and to have remained at about 1 percent (MacDorman et al., 2010).

Each decade, the methods for interpolating and extrapolating the BRT factors have been modified, but the initial factors have been altered only once. Research conducted during the 1980s revealed a "distinct and anomalous cohort effect" in the differences between the DA estimates and census counts for the Black population born between 1935 and 1945. The conclusion was that the 1940 Black BRT factors were over estimating the number of Black births that were not registered. Passel (1992) used regression methods to revise the series of registration completeness factors for Black births from 1935 to 1950. The corrected births based on these factors were incorporated into the 1990 DA estimates, and led to a downward revision to the number of Black births (1935 to 1950). This revision is also supported by an analysis conducted by Preston, Elo, Foster, and Fu (1998) and is included in the 2010 DA estimates. Additional information on the revisions to the extrapolations and interpolations is available from Robinson (2010).

To begin to understand the uncertainty around the estimated number of births used in DA, a high and low series were developed from an analysis prepared for DA in 1991 (Robinson, 1991). In 2010, the results of the 1991 work are considered still applicable because the birth registration tests that the correction factors were based on for 1945 through 1985 remained the same. The 1991 work examined four sources of uncertainty in the BRT factors (matching bias, correlation bias, interpolation/extrapolation error, and sampling variance) to develop an overall estimate of the possible uncertainty in the estimates of the completeness of birth registration. Results from each of these models were then grouped to develop lower and upper bound combined error factors. The lower bound factors primarily reflect the possible error due to matching bias, while the upper bound factors primarily reflect the possible error due to correlation bias. The ranges established in this report were carried forward and used to develop a low and high estimate of births occurring from 1945 through 1984. The low and high series of births were then compared with Census 2000 counts of the native-born population and the Medicare-based DA estimates for ages 65 to 74.

Without direct measures of the size of each source of error, professional judgment played a large role in specifying the models that were used to establish a range of uncertainty around the corrected births in the 1991 analysis. The intent was to establish a basis for understanding the range around the estimates of birth registration completeness from the BRTs under clearly-

specified scenarios, leaving researchers with the ability to incorporate their own assumptions and possibly arrive at a different result. A detailed description of these models is available in the full 1991 analysis.

The application of the lower bound factors from the 1991 analysis results in an estimated 249.0 million births between 1945 and 2010. The application of the higher bound factors from the 1991 analysis results in an estimated 252.4 million births between 1945 and 2010. These estimates can be compared with the estimated 249.9 million births that result from using the measures from the BRT factors alone, and 248.3 million when no corrections are made for completeness. The use of registered births with no correction factors applied as a lower bound estimate was ruled out because it was deemed demographically implausible. Because the range of uncertainty is developed around the number of births added based on the correction for completeness, the range is larger in earlier years where larger corrections are made, and zero in the years after 1984 when births registration is assumed to be complete.

Two comparisons were made to determine if the high and low series of births based on the 1991 work were plausible: comparisons were made between the estimated births and the Census 2000 counts of the native-born population by birth year, and births occurring between 1935 and 1945 were compared to the Medicare-based estimates for the corresponding estimate by age. The comparison suggested that the high series of births was implausible. It also showed that the corrected births currently being used were more consistent with both the Census 2000 native-born counts and the Medicare-based estimates than the high and low estimated births based on the factors from the 1991 report. However, there is still uncertainty around the number of births used in the DA estimates, and it remains likely that the actual number of births could be higher or lower than corrected births. In order to determine the number of high and low births to use in the range of DA estimates, several series of estimated births were developed using various levels of corrected births. The different series were obtained by multiplying the estimated number of unregistered births by age, sex, and race by 0.25, 0.30, 0.50, and 0.75. The results, along with the initial high and low values, were then also examined relative to the Census 2000 count of the native born and the Medicare-based estimates for ages 65 to 74. Using the 0.3 factor to increase and decrease the number of births added due to completeness results in a range of DA estimates

that comes closest to encompassing the Medicare-based estimates for each age 65 to 74. This adjustment also created a range of estimated births that was plausible when compared to the Census 2000 count of the native born. The DA plausible low series includes births that are based on the 0.3 downward alteration, and the plausible high series includes the 0.3 upward alteration. The result is three estimates of births that occurred between April 1, 1945 and April 1, 2010; a low estimate of 249.4 million births, a middle estimate of 249.9 million births, and a high estimate of 250.4 million births. These alterations were included mainly to demonstrate the impact of varying the level of the births added to account for underregistration. Different assumptions about the completeness of birth registration after 1985, when registration is assumed to be complete, were not used to develop alternative birth series. However, it is fairly easy to consider the impact of assuming different levels of completeness for these births after 1985 in that any assumption about completeness would have a corresponding increase in the size of that birth cohort. Table 2 identifies the series of births that was used in each of the five DA estimate series. Total births for 1935 to 2010 for each of the DA birth series, registered births used in DA, and NCHS published births are provided in Table 3.

Deaths

Uncertainty in the DA estimates due to the completeness of death registration has received little attention because it is not thought to be a large factor in the uncertainty surrounding the DA estimates. For use in DA, it is assumed that non-infant death registration has been complete since 1935. The use of Medicare data to estimate the population 65 and over minimizes the impact of potential inaccuracies in the recording of deaths or the recorded age at time of death on the DA estimates because the majority of deaths in the historical death data will have occurred to people who would have been older than age 65 on April 1, 2010.

Infant deaths are assumed to have been registered at half the race-specific level of the incompleteness of births up to and including 1959, and to have been completely registered beginning in 1960. An analysis of alternative assumptions of infant deaths showed that the impact is small overall. If it is assumed that infant deaths are under-reported at the same rate as births rather than half, the number of deaths in 1935, the year when the impact would be expected to be the largest, is increased by less than 8,000. However, aside from potential inclusion in the birth

registration completeness factors, if a birth was not registered and the corresponding infant death was also not registered, there would be no impact on the accuracy of the DA estimate.

Because of the lack of support for alternative assumptions about the number of deaths, only one series of deaths was used in the five DA series of estimates. All of the DA estimates include an estimated 14.8 million deaths to those who were born on or after April 1, 1945.

Projected values

At the time of the development of the DA estimates for this paper, final data from NCHS were not available for the period from January of 2008 through March of 2010. Therefore, a combination of projections, provisional data, and preliminary data from NCHS were used.[10] First, in order to estimate 2008 births, preliminary NCHS totals for 2008 are parsed out by month based on 2008 provisional monthly totals. Characteristics of the 2007 final births by month are then imposed on the 2008 monthly totals. Provisional 2009 NCHS monthly totals were used to estimate 2009 births, and again, the distribution of the characteristics of the 2007 births were imposed. Finally, for January 2010 through March 2010, values for the corresponding months in 2009 were used.

Missing values

To make full use of the NCHS vital statistics data, missing values for key variables needed to be assigned. The procedures used to obtain a value for a missing race of mother, race of father, or age at death are described below.

Race of Mother

Prior to 1988, NCHS did not impute the race of the mother if it was missing. For the years 1980 to 1988 when micro-level birth records were used, if the race of the mother was missing, the last known value after a sort on county of residence and year was used to impute the missing value. Between 1980 and 1988, about 88 thousand records out of over 30 million were missing race of mother.

[10] Birth and death data are released in three stages by NCHS: provisional, preliminary, and final. Data in each successive stage are more complete than in the previous stage. Provisional data are based on counts of events and may have incomplete medical and demographic information. The DA estimates will be revised once the final NCHS data are available.

Race of Father

From 1980 through 2007, the race of the father is missing on about 16 million records, or about 15 percent of almost 111 million records. When the father's race was missing, the mother's race was used for the race of the father.

Month and Year of Birth for Deaths

For each death, the month and year of birth are needed so that they can be subtracted from the correct birth cohort. If both month and year of birth are missing, it is possible to use the age at death and the month and year of death to determine the year of birth between two years. This is the case for 0.04 percent of the deaths from 2000 through 2007. For the same period, only 0.01 percent of deaths were missing both year of birth and age at death. This value is higher for earlier years. For the period between 1990 and 1999, 0.02 percent of deaths were missing both year of birth and age at death. This value for the period between 1980 and 1989 was 0.03 percent. To obtain the month and year of birth for the 1990 through 2007 records, if the month of birth is missing, we first set it to equal the month of death. If year of birth is missing, it is obtained by using the age at death, month of death, and either the reported or assigned month of birth. If both the age at death and year of birth are unknown, then month and year of birth are imputed from the last known value. For the 1980 through 1989 records, because of the large number of records with an age at time of death and a date of death, but a missing year and month of birth (89.6 percent), deaths were distributed between the two possible birth cohorts.

Uncertainty in the international migration component

International migration is the second largest component of population change in the DA estimates. Over the past decade the Census Bureau has undertaken a major initiative to improve its measures of net international migration. This has occurred primarily through the use of data from the ACS, which was not available in previous decades. To assess the degree of uncertainty in the estimates of international migration this decade, a range of estimates was developed that incorporated both results from different approaches for estimating each of the components of international migration and statistical measures of uncertainty where applicable. The DA estimate of net international migration reflects estimates for several distinct components of international

migration that are developed individually using different techniques. These components include: foreign-born immigration, foreign-born emigration, net migration between the United States and Puerto Rico, and the net migration of the native born. Each of the following sections focuses on a specific component, providing a description of how it was estimated as well as the process of accounting for the uncertainty around the component.

Foreign-born immigration

Estimates of foreign-born immigration (the number of those who were foreign born who entered the United States) were based on two questions asked in the ACS. First, estimates of foreign-born immigration were developed using the residence one year ago (ROYA) question. The foreign born who reported in the ACS that their residence one year ago was "abroad" are considered immigrants. Because this question is asked only to those aged one and higher, an additional assumption was made to account for immigrants under the age of one. It was assumed that the number of immigrants under the age of one was half the number of those age one whose residence one year ago was abroad. As an alternative to the ROYA-based estimate, data collected in the ACS on year of entry (YOE) were used. With this approach, the foreign born whose year of entry is the year prior to the survey year are considered immigrants. The YOE method resulted in higher estimates of foreign-born immigration than the ROYA method.[11] The ACS-based estimates are developed for the period from July 1 to June 1. This estimate is divided by twelve to get monthly estimates that can be used to obtain estimates by census year.

When using the ROYA method to estimate foreign-born immigration, total immigration for all ages from April 1, 2000 to April 1, 2010 was estimated to be 11.9 million. When the estimate is based on YOE, the estimate rises to 13.5 million. Both the ROYA and YOE methods are developed from a sample of the population and are subject to sampling error. The 90-percent confidence interval around the ROYA and YOE estimates was used to develop additional estimates of foreign-born immigration. When the lower bound of the 90-percent confidence interval from the ROYA estimate is used, the estimated foreign-born immigration is 11.4

[11] No distinction is made between immigrants based on their legal status when using the ROYA and YOE questions on the ACS to estimate migration of the foreign born. The ACS does not include questions on the legal status of the foreign born.

million; when the upper bound of the 90-percent confidence interval YOE estimate is used; the estimated foreign-born immigration is 13.9 million.

Both the ROYA and YOE estimates used in the 2010 DA were developed using single-year data from the 2000 to 2009 ACS. For these time series, the 2010 value was calculated by holding the 2009 estimates constant. Since 2007, the annual estimates of foreign-born immigration have declined, and so by holding the 2009 estimate constant, foreign-born immigration for 2010 may be overstated. However, the impact of holding the 2009 ACS estimates constant to obtain the estimated total foreign-born immigration for the remainder of the decade is thought to be relatively small.

Foreign-born emigration

Estimates of foreign-born emigration were developed using a residual method to estimate rates of emigration which were then applied to estimates of the foreign-born population from the ACS. The foreign-born household population in Census 2000 is aged forward using NCHS life tables to obtain the expected population in 2005, 2006, 2007, 2008, and 2009. The expected foreign-born population estimates for those years are then compared to the foreign-born population estimated by the ACS. Subtracting the estimated from the expected population produces a residual, which serves as the basis for emigration rates. This calculation is performed for two period-of-entry groups: the foreign born who entered the United States between 1990 and 1999; and the foreign born who entered before 1990. An average of the rates for each period-of-entry group is then applied to the population at risk of emigrating each year (i.e., the foreign-born population in the ACS who indicated that they lived in the United States one year ago) to obtain an annual estimate of emigration for 2000 to 2009. Using this method, we estimate 2.3 million foreign-born emigrants for the April 1, 2000 to April 1, 2010 time period.

In order to evaluate the uncertainty around this estimate due to sampling error, the 90-percent confidence intervals around the ACS estimates were used to develop high and low estimates of emigration. To develop a low estimate of foreign-born emigration, the upper bound of the ACS estimate in the residual calculation is used to develop an estimated low rate of emigration.

These rates are then applied to the lower bound estimates of the foreign-born population in each year of the ACS to determine the low annual estimate of foreign-born emigration. Similarly, to develop a high estimate of foreign-born emigration, the lower bound of the ACS estimate in the residual calculation is used to develop a high rate of emigration. This high rate is then applied to the upper bound estimates from the ACS to develop a high estimate of foreign-born emigration. These calculations result in a high emigration estimate of 2.6 million and a low emigration estimate of 2.0 million.

Net native-born migration

Estimates of net native-born migration were based primarily on work by Schachter (2008) that examined census data from other countries. The estimates were developed using a residual method. The number of those native to the United States was ascertained from a census from each county, aged forward, and then compared to the number of those native to the United States in a later census from the same country. The results of that work indicated an annual net loss due to native-born migration of just over 45 thousand people. One issue identified with the approach used by Schachter is that what countries measure in their census can vary. Some countries (e.g., Germany) define nativity as citizenship in their census. This means that even those who are born in Germany may not be considered German citizens if their parents are not citizens. In other countries, only data by place of birth were available; therefore, children born in these countries of U.S. citizen parents would not be identified as a U.S. citizen. This difference poses problems for the coverage of the population that is born abroad of U.S. citizen parents in Schachter's net native-born migration measure. Therefore, three different assumptions were made about the role of those born abroad of U.S. citizen parents in net native-born migration. First, it was assumed that all those born abroad of U.S. citizen parents are included in the measure of net native-born migration, resulting in the annual estimate of negative 45 thousand. Second, it was assumed that none of the population born abroad of U.S. citizen parents were included in the net native-born migration measure. To estimate this population, the Census 2000 born abroad of U.S. citizen parent enumeration was survived forward using NCHS death rates and compared to the born abroad population enumeration in the ACS 2009. The difference, after projecting to April 1, 2010, results in an estimate of the net inflow of this population over the decade of approximately 624 thousand. Third, it was assumed that those

born abroad to U.S. citizen parents in countries that only measure citizenship in their census would be included in Schachter's estimate (and therefore included in the estimate of net native-born migration) while those born in other countries would not. If those countries who only measure citizenship are not included, the result is an estimated net inflow of this population of approximately 397 thousand over the decade. Data collected and provided by the Defense Manpower Data Center are used to estimate the stock of the Armed Forces population living overseas as of April 1, 2010. The stock of the AFO population is subtracted from the estimate of the resident population.

Net Migration between the United States and Puerto Rico

Net migration between the United States and Puerto Rico is also estimated using data from the ACS, and also includes data from the Puerto Rico Community Survey (PRCS) for 2005 and later. People who indicated in the ACS that they lived in Puerto Rico one year ago are considered immigrants (i.e., they moved from Puerto Rico to one of the 50 states or the District of Columbia). People who indicated on the PRCS that they lived in the United States one year ago are considered emigrants (i.e., they moved from one of the 50 states or the District of Columbia to Puerto Rico). We assume the number of immigrants and emigrants under the age of one is equal to half of the number of one-year-old immigrants and emigrants, respectively. This calculation results in an estimate of approximately 205 thousand net migrants from Puerto Rico to the United States over the decade.

Modification of the 2000 DA estimate of the foreign-born population

The 2000 DA estimate of net international migration utilized a residual method to estimate undocumented migration. This approach included the assumption that the Census 2000 count of the "residual" component of the foreign born was only 85 percent complete. For the DA plausible low series, this assumption was replaced with a coverage profile based on information from the A.C.E. Revision II and CPS coverage rates. Specifically, we compared the Dual System Estimates (DSE) from the A.C.E. Revision II to estimates from Census 2000 to identify patterns in coverage by age, sex, and Hispanic origin. We also evaluated coverage patterns in the CPS by age, sex, and Hispanic origin (U.S. Census Bureau, 2009). The values used in the coverage profile are provided in Table 4. This modification resulted in a reduction to the 2000

DA foreign-born population of just over 1 million people. More information on the development of the 2000 DA estimate of net international migration is available in Appendix A of the ESCAP II: Demographic Analysis Results report (U.S. Census Bureau, 2001).

Accounting for representation of the foreign born in the ACS

Foreign-born immigration from 2000 to 2010 is estimated primarily using data from the ACS. The representation of recent foreign-born immigrants in the ACS is unknown, and underrepresentation may result in an underestimate of foreign-born immigration. Currently, with both the ROYA- and YOE-based immigration estimates, we assume that underrepresentation of foreign-born immigrants in the ACS is accounted for through the application of survey weights and population controls. However, the ACS is controlled by age, sex, race, and Hispanic origin, but not nativity, and so these controls might not account entirely for the underrepresentation of the foreign-born population. To assess the potential impact of underrepresentation of the foreign-born population in the ACS, a series of representation factors were developed and applied to single-year ACS data to produce a new series of foreign-born immigration estimates.

The representation factors were developed by comparing pre-population controlled data from the Census 2000 Supplementary Survey (C2SS) to data from Census 2000 by broad demographic characteristics, including nativity. Specifically, we divided estimates of the foreign-born population by broad age group, sex, and Hispanic origin from Census 2000 by estimates of the foreign-born population in the C2SS. The resulting representation factors were then applied to single year data from the 2000 to 2009 ACS to provide estimates of foreign-born immigration.[12] In addition, we developed a series of factors that attempt to account for potential sampling error in the C2SS by subtracting the margin of error from the population estimate in each cell and then calculating the coverage factors using the same methodology discussed above. Applying these factors to YOE-based estimates produces a high estimate of foreign-born immigration of 14.2 million over the decade.

[12] The 2000 C2SS is different than the 2000 ACS. The C2SS was an expansion of the ACS and is more comparable to later years of the ACS than the 2000 ACS.

The accuracy of the representation factors relies on multiple assumptions: 1) coverage of Census 2000 was 100 percent; 2) representation factors developed using data from the 2000 C2SS are applicable to later years of the ACS; 3) representation of the total foreign-born population is the same as the representation of recent immigrants; and 4) there is no variation in coverage within each cell used to calculate the representation factors.

Alternative estimates of international migration

In addition to the above methods which were included in the DA estimates, we evaluated results developed using alternative approaches to further assess the estimates.

As one set of alternative estimates, two "change in stock" measures were calculated to estimate net foreign-born migration during the decade. The first change in stock measure was calculated beginning with the foreign-born stock in Census 2000. This population was survived forward to July 1, 2009 using the cohort-component method, resulting in an estimated 2.3 million deaths or an estimated foreign-born population of 28.9 million. This number was compared to the foreign-born population from the single-year 2009 ACS file of 38.5 million. This results in an implied growth in the foreign-born population from April 1, 2000 to July 1, 2009 of 9.7 million. This population growth was then projected forward to April 1, 2010, resulting in an estimated growth of the foreign-born population over the decade of 10.5 million.

A second estimate involved using administrative, Census 2000, and ACS data to first develop an estimate of the foreign-born stock in 2010, and then perform a similar calculation to that described above to estimate net foreign-born migration from 2000 to 2010. This estimate started with the native-born population in Census 2000, which was then survived forward to April 1, 2010 using the cohort component method with birth and death rates from NCHS.[13] The same assumptions about net native migration used in the ROYA and YOE estimates were also used in these calculations, resulting in a yearly decrease of approximately 45 thousand in the native-born population (Schachter, 2008). This results in an estimate of 269.0 million native born on April 1, 2010. A ratio of the foreign-born population to the native-born population by single year of age and sex from the controlled three-year ACS file (2006-2008) is then applied to the April 1, 2010

[13] NCHS life tables were available for years up to 2006. After 2006, death rates are held constant.

native-born estimate. The resulting estimate of the foreign born is 38.3 million. The foreign-born population in Census 2000 survived forward using NCHS death rates is then used to ascertain the implied increase between Census 2000 and April 1, 2010. The Census 2000 count of foreign born was 31.1 million with an estimated 2.5 million deaths occurring over the decade, resulting in an implied increase in the foreign-born population of 9.7 million. The same calculations can be carried out using ratios from the ACS data without the population controls applied. Using three-year ACS data (2006-2008) without the population controls applied to calculate the foreign born to native-born ratio the implied increase in the foreign-born population over the decade is 10.0 million (an increase of 0.3 million compared to using the final weighted ACS data).

Range of estimated net international migration

The range of estimates of net international migration from Apri1 1, 2000 to April 1, 2010 used in the DA estimates for all ages included five estimates that ranged from a low of 8.6 million to a high of 12.6 million. The lowest estimate of international migration, which was used in the plausible low series for the population under age 65, used the lower 90-percent confidence interval ROYA estimate and the high estimate of foreign-born emigration. The plausible low series also includes a reduction to the 2000 DA foreign-born population of just over one million people, but this does not impact the estimates of net international migration for this period. The highest estimate, which was used in the plausible high series, uses the YOE estimate with the ACS coverage factors applied, and the low estimate of foreign-born emigration. The plausible high series does not include the reduction to the 2000 DA foreign-born population. Annual estimates for 2000 through 2010 for each component of international migration are provided in Table 5. Table 2 lists which estimate of each component was used in the each of the five DA estimate series.

Uncertainty in the Medicare-based estimates of the population 65 and over

The DA estimates for the population aged 65 and over were developed from tabulations of those enrolled in Medicare in 2009 and estimates of those not enrolled.[14] The estimates of those not

[14] Enrollees are aged forward one year to estimate those enrolled in 2010. The DA estimates will be revised using a 2010 Medicare enrollment file when it is available.

enrolled were developed by combining information on the timing of enrollment included on the Medicare file with data from the CPS question on health care coverage. The accuracy of the Medicare-based estimates depends on the ability to define the proper universe of Medicare enrollees, the ability to accurately identify and remove duplicate and inaccurate records from the Medicare enrollment database (MEDB), and the accuracy of the estimate of the number of those not enrolled.

In general, U.S. citizens or permanent residents are eligible for Medicare benefits if they have worked for at least ten years in a job that has paid money into the Medicare system. The eligibility rule also applies to spouses. If either spouse paid money into the Medicare system, then they are both eligible. Those who have worked less than the ten years may also be eligible but will have additional costs. Legal immigrants must have had continuous residency in the United States for at least five years to be eligible to enroll.

Because of questions about the completeness of the historical vital statistics data for years prior to 1935, past DA efforts have relied primarily on the Medicare data to estimate the population 65 and over. For the first time, there will be an overlap between the Medicare-based estimates and the estimates developed using the 1935 through 1944 vital statistics data and estimates of net international migration. The overlap includes the population 65 to 74 years of age. This overlap, along with a comparison between estimates of those not enrolled from the CPS and estimates based on the recently-added question on health care coverage in the ACS, forms the basis for understanding the uncertainty in the Medicare-based estimates.

Removal of records from the MEDB

The MEDB is maintained by the Centers for Medicare and Medicaid Services. Several types of records were excluded from the MEDB to obtain a dataset of records believed to best match the population 65 and older on April 1, 2010 who met the 2010 Census residency rules. Records were excluded when the enrollee: did not reside in the United States; was not aged 65 or older as of April 1; had no date of birth; or had a date of death before April 1, 2010.[15] In addition,

[15] Some individuals under age 65 may be eligible for Medicare if diagnosed with permanent disability or permanent kidney failure that requires dialysis or a kidney transplant.

records for those with an implied age of 115 or older were removed and attempts were made to remove duplicate records.

In order to assign race to the records without a race classification including records classified as only Hispanic, the Medicare file was merged to the Census Bureau's Person Characteristics File (PCF). After the merge, records that were not unique (i.e., they were duplicates of other records) were removed. Overall, approximately 38 thousand records (about 0.10 percent of the original universe) were removed. The female population lost more records than the male population. About 32 thousand records, or 0.17 percent, were removed from the non-Black females and close to 2,700, or 0.14 percent, from the Black females.[16]

Many records show an age of 115 or older. More than likely, the presence of this high number of extreme ages represents erroneous inclusions on the file of deceased individuals. Age is calculated from information about date of birth, and proof of age is necessary to enroll in Medicare. It is therefore unlikely that there would be intentional age misreporting. Instead, date of birth (year) may have been entered into the database incorrectly. After the unduplication process, about 41 thousand records with an implied age of 115 or older were removed from the file.

Ultimately, about 79 thousand (or 0.21 percent) of the total records were removed from the Medicare universe because they were thought to represent either duplicate or erroneous records.

Estimates of the number of people not enrolled

Measures from different sources indicate that Medicare enrollment is a nearly complete source of information on the size of the population aged 65 and older. In the late 1990s, the Social Security Administration estimated the Medicare coverage of the population to be about 96 percent (U.S. House Committee on Ways and Means, 1998).

[16] Through further inspection of the file, an additional 1,607 pairs of records were identified that had identical protected identification codes (PIKs), age (identical date of birth), and sex, but had different race codes. These records also appeared to be duplicates. However, in the absence of knowing which record in a pair to keep, these records were not removed from the file.

The accuracy of population estimates produced from the MEDB depends on the accuracy of the estimated completeness of the database. Some groups are not eligible to enroll, some may delay enrollment until a date later than when they became eligible; and some may never enroll. Groups not eligible to enroll include federal workers covered under a former federal employee retirement program, the Civil Service Retirement System (CSRS), annuitants without the required quarters of Social Security coverage, and noncitizens who have been residents for less than five years.[17]

The DA methodology has traditionally developed and combined two factors to correct the database for those not enrolled: an estimate of those who delay enrolling past their initial age of eligibility, and an estimate of those who will never enroll. The two factors are combined to form a total underenrollment correction, which is then applied to the database by single year of age, sex, and race (Black and non-Black).

The calculation of the delayed enrollment factor was based on information on age at time of enrollment in the MEDB for those aged 72 to 84. To get a sufficient number of years of "exposure" to delayed enrollment, ages 65 to 72 were excluded from the derivation. It was assumed that no one delays enrollment beyond age 85. The validity of these assumptions was based on a cohort analysis of Medicare enrollment in the 1980s (Passel and Robinson, 1988). Since the estimates of delayed enrollment are based on information from those currently enrolled within a specified age range, the estimates are sensitive to the actual changes in enrollment patterns that may have occurred.

The delayed enrollment adjustments relate to the population that will eventually enroll in Medicare at some age. In order to estimate the total population not enrolled, we also need an estimate of those who will never enroll. Previous research has assumed that individuals who delay enrollment will have enrolled by age 85. Thus, the percent of the population aged 85 and over that is not enrolled in Medicare forms the basis for the estimate of the population aged 65 and over who will never enroll in Medicare. The CPS Annual Social and Economic Supplement

[17] The Civil Service Retirement System was a previous federal employee retirement program that stopped adding new participants during the 1980s.

(ASEC) question on health insurance serves as the source for the development of the never-enrolled factor.[18] The CPS ASEC includes a specific question about enrollment in Medicare.

The percentages for both the delayed and the never-enrolled factors are added to derive the total percent not enrolled specific to each combination of single year of age, sex, and race (Black and non-Black). The factor is then applied to the MEDB enrollment number (by age, sex, and race) to produce the DA estimates for ages 65 and over.

Two basic approaches were used to estimate the number of those not enrolled in Medicare. The first approach (series 1) uses factors for those who will delay enrollment developed from the 2009 MEDB and factors for those who will never enroll developed from estimates of enrollment from the CPS for 2002-2008 for the retired population aged 75 and older. The second approach (series 2) uses estimates of enrollment from the CPS for the total population in key age groups to estimate the number of those who are either delaying enrollment or who will never enroll. From these two approaches, alternative enrollment factors were also developed to show the potential impact of sampling error. While not used in the five series of DA estimates, underenrollment factors were also developed using data from the ACS. Lastly, as a check on both the Medicare- and component-based estimates, underenrollment factors were developed from the differences between the number of records in the MEDB and the component-based estimates for those aged 65 to 74 (the Medicare-component overlap).

In 2000 DA, the CPS estimates of underenrollment for the population aged 85 and older were used to determine the percent that will never enroll in Medicare beyond age 85. The number of those who will never enroll must also be calculated for those aged 65 to 84. For non-Black males and non-Black females, the percentages for the 85 and over population in 2000 were kept constant across all age groups 65 to 84 (2.0 and 1.5 percent, respectively).

However, it was assumed that the enrollment levels for Black males and females aged 65 to 69 had become more complete than for those over 85 years of age, unlike their non-Black

[18] The Current Population Survey (CPS) is a monthly survey of about 50 thousand households conducted by the Census Bureau for the Bureau of Labor Statistics. The survey has been conducted for more than 60 years. The data are collected through a combination of telephone and in-person modes using computer-assisted instruments.

counterparts. For the Black male population, the percentages for those who will never enroll were assumed to decline from 3.5 percent for those aged 85 and over to 2.5 percent for those aged 65 to 69. For the Black female population, the percentages were assumed to decline from 2.5 percent to 2.0 percent. The factors for those who will never enroll for the 70 to 84 age group were obtained through interpolation.

Medicare Series 1

For the first Medicare series, the methodology for the delayed enrollment factors remained the same as in 2000, but was updated using the 2009 MEDB file.

The CPS enrollment estimates for ages 75 and older were used to set the percent that will never enroll in Medicare at ages 85 and older. Because they were thought to better represent the population 85 and older where almost all persons are retired, the Medicare enrollment estimates for the retired population were used instead of the estimates for the total population (including the retired and those in the labor force). The never-enrolled factors are then held at the age 75 and older level for younger ages for both the Black and non-Black population. The first series of Medicare-based estimates were used as the middle DA estimates for the population aged 65 and over.

Because the CPS estimates are based on sample surveys, additional estimates were developed to show the potential impact of sampling error. As with the estimates of international migration, the 90-percent confidence interval was used. The estimates based on the lower bound of the 90-percent confidence interval were used for the low estimate of the population aged 65 and over.

Medicare Series 2

For DA in 2000, and for the first Medicare series, the measurement of those who delay enrollment and those who will never enroll were treated as two different measurements; one factor is perceived to be capturing those who will *eventually* enroll while the other factor is perceived to compensate for those who will *never* enroll. Though conceptually distinct, it is possible that there is an overlap in what is being measured. It is also possible that neither factor captures the entire population.

For the second series of Medicare-based estimates, it was assumed that the correction factor derived directly from the CPS captures the pace of enrollment as well as the proportion of the population who will never enroll. The factors are calculated by sex, race, and 5-year age groups for ages 65 to 69 and 70 to 74. It is assumed that the enrollment pattern is constant after age 75. The factors for the second series include all respondents in each age group, rather than just the retired population. The second series of Medicare-based estimates was used for the high DA estimate. Table 2 identifies which Medicare-based estimate was used in the each of the five DA estimate series.

Table 6 shows the estimates produced using the different approaches as well as the upper and lower bounds from the margins of error. The total population ranges from a low of 39.8 million when using the correction factors developed for the 1990s and DA in 2000 to a high of almost 40.7 million when the factors based on CPS data collected in the 2002-2008 period are applied directly to the Medicare enrollment database. When considering sampling variability in the correction factors, the upper value of the range increases to 40.9 million.

The CPS universe (civilian noninstitutionalized population only)

The CPS is only administered to the civilian noninstitutionalized population, i.e., the institutionalized population is not included in the universe. This might affect the estimate of persons not enrolled in Medicare. However, we consider the impact to be small. The implicit assumption is that the institutionalized population is not enrolled in Medicare at the same rate as the noninstitutionalized population. Overall, according to Census 2000 data, 1.6 million persons were institutionalized at ages 65 and older. Calculations based on the estimates of the resident population and estimates of the noninstitutionalized population by age and sex show that the institutionalized population does not become a significant portion of the population until it reaches age 85.[19] In this age category (ages 85 and older), around 21 percent of females and about 19 percent of males are in institutions. Furthermore, a large portion of the institutionalized population is likely enrolled in Medicare.

[19] Special tabulations produced by the Population Division's Estimates and Projections Program.

Comparison with the component-based estimates (the age 65 to 74 overlap)

For DA in 2010, the primary estimate for the population aged 65 and older was developed from aggregated Medicare data corrected for underenrollment. Estimates were also developed for the population aged 65 to 74 using the component approach. The availability of the second estimate—the ages 65 to 74 overlap—allows us to assess the consistency between the Medicare- and component-based estimates for these ages by sex and Black and non-Black.

The 2000 DA population aged 56 to 64 in 2000 reached ages 65 to 73 in 2009. We carried forward the 2000 DA estimates through the cohort-component approach to form a 2009 DA estimate. This estimate was compared to the Medicare 2009 enrollment universe (before corrections for underenrollment and projections to 2010). The difference was interpreted to represent an indirect demographic estimate of the MEDB completeness on April 1, 2009.

The patterns of difference in the implied underenrollment vary by race more than by sex. For the non-Black population aged 70 to 73 in 2009, the implied underenrollment percents based on the component DA estimates (4.0 for males, 2.7 for females) are reasonably similar to the Medicare/CPS-based percentages for ages 70 to 74 (3.3 to 4.4 for males, 2.6 to 4.0 for females). Given the different estimation methods, this agreement is notable. The DA estimates are based on births from 1935 to 1939, corrected for underregistration—these birth cohorts are then carried forward 70 or more years by subtracting deaths and adding estimates of net international migration. The Medicare-based estimates are derived from factors from the 2009 Medicare file and CPS estimates of Medicare underenrollment for 2002-2008.

However, for the non-Black population aged 65 to 69, the component-based estimates of underenrollment (4.5 percent for males, 3.4 percent for females) are lower than the Medicare/CPS-based estimates, especially the series that was based directly on the CPS estimates of Medicare coverage, which imply relatively high levels of underenrollment (11.2 percent for males, 9.6 percent for females). We need to investigate reasons for the difference, and see if these reasons can also explain why the two sets are so similar for ages 70 to 73.

For the Black population, the age pattern of agreement is different. This time, the component-based estimates of implied underenrollment at ages 65 to 69 (12.0 for males, 8.2 for females) are

broadly similar to the Medicare/CPS-based estimates, in that all show high percentages of Medicare underenrollment. The estimates based entirely on the CPS again tend to be on the high side, especially for Black females aged 65 to 69 (12.1 percent). For Black females aged 70 to 73 in 2009, the component-based implied percent (5.8) also falls within the range of the Medicare/CPS-based estimates (3.8 to 6.8). The distinct exception is for Black males aged 70 to 73, where the Medicare underenrollment of 11.0 percent implied by the component DA estimate is appreciably higher than every Medicare/CPS-based set (4.6 to 7.6). The availability of the 65 to 74 "overlap" provides a new opportunity to systematically evaluate both the Medicare-based estimates and the component-based estimates. The factors that were used to develop each series of Medicare-based estimates and the implied underenrollment based on the 65-74 overlap are provided in Table 7.

Estimates of underenrollment from the ACS

To provide additional support for the reasonableness of the CPS-based factors, we also looked at Medicare enrollment data from the ACS. A question on health care coverage that asks specifically about enrollment in Medicare was added to the ACS in 2008, and the first estimates from this question were released in 2009. For the population aged 65 and older, the estimated percent enrolled from the ACS was 96.9, and 93.4 percent from the CPS.

Differences in the CPS and the ACS questionnaire wording and collection mode might lead to different responses about health care enrollment. The 2008 ACS asks the health insurance question about each person individually (does this person have coverage). The CPS asks the question at the household level (does anyone in the household have coverage) and then asks who in the household had coverage. The different approaches might generate different outcomes. The use of different data collection modes and edits for nonresponse may also lead to different results.

OBTAINING THE DA ESTIMATES BY RACE (BLACK AND NON-BLACK)

The previous sections discussed each component of the DA estimates with a focus on the estimates of the total population. The following section describes how the estimates by Black and non-Black are developed from each component.

Birth records

Births recorded through birth certificates serve as the primary source for determining the size of each native-born cohort. Birth certificates do not include information on the race and Hispanic origin of the child, and so it must be inferred from the race and ethnicity of the parents. The information recorded on the birth certificate on the race(s) and Hispanic origin of the mother and father are used to assign a race and ethnicity to each birth. Decisions about how to use this information and how the decennial census responses are tabulated must be considered when comparisons are being made between the DA estimates and the 2010 Census counts.

Following guidance provided by OMB's 1997 revision to the federal standards on collecting information on race and ethnicity, NCHS implemented the new standards in 2003 to include the option of selecting more than one race on birth and death certificates, but not all states have adopted this standard. The OMB revision specified five minimum categories for data on race and two for data on ethnicity: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White for race, and "Hispanic or Latino" and "Not Hispanic or Latino" for ethnicity. NCHS was given the option, which they exercised, to also include the category "Other." By 2006, 23 states had implemented a revised birth certificate to allow for the reporting of more than one race. The 23 states that reported multiple races for births accounted for 55 percent of U.S. births in 2006. Of the births reported in these states, 1.6 percent were to mothers who identified as multiracial (U.S. Department of Health and Human Services, 2008). NCHS provides both the multiple races that are reported and the multiple race responses "bridged" to the pre-1997 OMB four single-race categories.[20] Even when the multiple race option was not available, having the race of the mother and father on the birth certificate makes it possible

[20] The OMB revisions separated the formerly combined Asian and Pacific Islander category to form five racial categories: 1) White; 2) Black or African American; 3) American Indian and Alaska Native; 4) Asian; and 5) Native Hawaiian and Other Pacific Islander.

to identify some births to parents of different races. However, if either parent was multiracial, they may not have had the opportunity to indicate so when their child was born, but will have the opportunity to identify themselves or their children as multiracial in the 2010 Census. An example would be a birth where the mother was Black (Black alone) and the father was Black and Asian (Black in combination). The birth may be recorded as being to two Black parents (Black alone) but reported as Black and Asian in the census (Black in combination). The value of the DA estimates by race as an indicator of the quality of the census depends largely on how effective we have been at obtaining consistency between the DA estimate and the census categories with which they are being compared.

For DA in 2000, race was assigned to births based on the race of the father because it produced an estimate of the Black population that had the closest agreement with the tabulated census responses from among other approaches based on the reported race of the parents (Adlahka et al., 2002). The Census 2000 value that compared to the DA estimate was an average of the count of the Black alone population and the Black alone or in combination population.

With the option of selecting multiple races in Census 2000 and in the 2010 Census, we need to consider how the race of each parent is recorded on the birth certificate and how this relates to the reporting of multiple races in the census. With regard to Hispanic origin, the issue is more similar to what was encountered in previous decades when only a single race was reported in the census but the race of both parents was available from the birth records: when information on Hispanic origin is available, the birth certificate will include the Hispanic origin of both parents, while in the census each respondent indicates if they are of Hispanic origin or not of Hispanic origin.

Since 2000, there has been an increase in both multiracial and ethnic births, as well as possible changes in how people self-identify in the census. Subsequently, for 2010 more effort has been directed toward researching how to best align the decennial census counts with the DA estimates for purposes of comparison. For DA in 2010, estimates were developed for the Black alone population and the Black alone or in combination population. This should allow for a direct comparison with Census 2010 tabulations where the Black population is defined as only those who

identified as Black alone, and tabulations where the Black population is defined as including those who identified as Black alone or Black in combination with another race.

The use of micro-level vital-statistics data in 2010 has allowed for greater flexibility in assigning race to births since 2000, while also making it possible to re-tabulate the birth data for earlier years. For years prior to 1980, the birth and death data used consist of annual tabulations provided by NCHS. For 1968 through 1979, we have estimates of births by race of father, race of mother, and race of child. For years before 1968, only tabulations by "race of child" as defined based on the NCHS "Minority rule" are available.[21] Because we only have tabulated data, there are limits to what we can do for births occurring in these years. However, because multiracial births were relatively rare prior to 1980, this does not have a large impact on the DA estimates by race. Also, this will not impact the estimates by Hispanic origin since those under age 20 in 2010 were born after 1980. The time series of births based on race of father was extended back to 1935 using the ratio estimated births by race using the race of father and the Minority rule.

The basic strategy for assigning the race categories needed for DA (Black alone, Black in combination, and not Black alone or in combination) was to identify a birth as non-Black if both parents on the birth certificate were non-Black, and Black if both parents were Black. If the parents' races on the birth certificate reflected a Black and non-Black combination, proportions developed from Census 2000 responses were used. To identify patterns in how the race of children in interracial/ethnic households were reported, a dataset was developed based on individual-level Census 2000 records which linked children aged 0 to 17 and their parents. This "Kid Link File" consists of four sets of variables: characteristics of the child; characteristics of the mother; characteristics of the father; and characteristics of the household. This format allows the reported characteristics of the child—such as race and Hispanic origin—to be linked to parents' reporting.[22]

[21] The Minority Rule, as used by NCHS determines the race for each birth as follows: when of the same race, the race of the birth is the same as the race of the parents; when of different races and one parent is White, the birth is assigned the race of the minority parent; when the parents are of different races and neither parent is White, the birth is assigned the father's race, unless either parent is Hawaiian; then the birth is assigned Hawaiian.
If the race was missing from one parent, the birth is assigned the race of the other parent.
[22] After results are available from the 2010 Census, the Kid Link file will be updated using 2010 Census responses.

The Kid Link file was used to calculate the proportions of children identified in Census 2000 as Black alone, Black in combination, or not Black alone or in combination according to the combination of parents' races. The householder, spouse of the householder, or unmarried partner of the householder were considered the parents. Because we wanted to best match the relationship between parents' race on the birth certificate and race reported in the census, households where the spouse or partner was of the same sex as the householder were excluded, and only natural-born children of the householder were included. If the race of any member of the household was imputed or identified as Some Other Race in the census, the household was excluded unless the Some Other Race response was in combination with another race category. Proportions were generated for every mother/father combination of White alone, Black alone, Asian alone, American Indian or Alaska Native alone, Native Hawaiian or Other Pacific Islander alone, and for a combined Asian/Native Hawaiian or Other Pacific Islander parent category. The parent category of Asian combined with Native Hawaiian or Other Pacific Islander was used because the majority of births between 1980 and 2007 were recorded using birth certificates with the combined category (Asian and Pacific Islander) rather than the separate Asian and Native Hawaiian or Other Pacific Islander categories. The resulting proportions were then applied to each birth based on the reported race of both parents on the birth certificate to obtain a race distribution consistent with the patterns of reporting in Census 2000.

This approach allowed us to address that when births occur to one Black parent and one non-Black parent, some are given the race of the mother and some the race of the father, but the majority are identified as multiracial rather than relying solely on the reported race of either parent. The proportion of births identified as Black in combination ranged from 56.5 percent for the combined Asian /Native Hawaiian or Other Pacific Islander Mother and Black Father category to 16.4 percent for the Black Mother and American Indian or Alaska Native Father category. The proportions for all of the parental race combinations are provided in Table 8. When both parents are of the same race (Black or non-Black), all of the births were assigned the race of their parents. The assignment of ethnicity will be discussed in more detail in the "Estimates by Hispanic Origin" section.

The DA estimates by race that will be disseminated prior to the release of the 2010 Census counts will include estimates for the Black alone population. Additional estimates for the Black alone or in combination population will be released at a later date.[23]

Characteristics of international migrants

To obtain estimates of the age, sex, racial, and ethnic characteristics for each component of international migration from 2000 to 2010, "proxy" populations that are thought to best represent the characteristics of each component are used. The characteristics of these proxy populations from Census 2000 and the 2005-2007 three-year ACS file are used to estimate the age, sex, racial, and ethnic characteristics for each component. The characteristics of the foreign-born population who entered the United States within 5 years of the Census/survey year are used to estimate the characteristics of foreign-born immigrants. Emigration rates were calculated for two period of entry groups: the foreign born who entered the United States between 1990 and 1999; and the foreign born who entered before 1990. The characteristics of the foreign born who entered the United States within ten years of the Census/survey year were used to estimate the characteristics of emigrants who entered within ten years of the estimate year. The characteristics of the foreign born who entered the United States more than ten years before the Census/survey year were used to estimate the characteristics of emigrants who entered the United States more than ten years before the estimate year. The characteristics of the population born in Puerto Rico who entered the United States within five/ten years of the Census/survey year are used to obtain the characteristics of the net migrants between the United States and Puerto Rico. The characteristics of the native population are used to obtain the characteristics of the net native migrants. Finally, the characteristics of the native population who were born abroad of U.S. citizen parents that entered the United States in the last 10 years were used to obtain the characteristics of those born abroad of U.S. citizen parents.

Within each component, the population that is Black alone was categorized as "Black" while the remaining population was categorized as "non-Black." Alternative classifications of race using the multiracial categories found in the 2000 Census and the ACS were also developed. For this

[23] The DA estimates for the Black population reflect the use of the Kid Link proportions for births occurring after April 1, 1980 and the use of the Father's race for prior years.

alternative classification, the Black alone and all multiple race groups that include Black were categorized as "Black alone or in combination," and the remaining population was categorized as "not Black alone or in combination." The alternative classifications were not used in the five series of DA estimates which only include estimates of the Black alone population, but will be used in future DA estimates for the Black alone or in combination population.

Characteristics of Medicare enrollees and those not enrolled

The Centers for Medicare and Medicaid (CMS) handles enrollment in the Medicare program. However, when the Medicare program began in 1965, the Social Security Administration (SSA) was given managerial responsibility and is responsible for certifying that an individual is eligible for Medicare and for transmitting demographic information about that individual to CMS. Race is one such piece of demographic information. It was obtained at the time of application to the SSA for a Social Security Number (SSN).

From 1936 (the beginning of Social Security) to 1980, the race data were collected in only three categories: White, Black, and Other with an additional category of Unknown race. Since 1980, SSA has expanded their race categories. On the SS-5 form (application for a Social Security card) the following race/ethnicity categories are listed: Asian, Asian-American[*sic*] or Pacific Islander, Hispanic, Black, North American Indian or Alaska Native and White (SIC).[24] The Medicare file has race distributed in these categories plus Other, Unknown (don't know or not ascertained) and Blank (missing). No attempts are made to recode race collected in the earlier categories into the revised categories. Race is a voluntary item on the Social Security card application. Unknown and Blank race classifications also stem from enrollees who did not come in through the regular application route for Social Security, such as former railroad workers who enroll through the Railroad Retirement Board or individuals who enroll through the Medicare Health Plan. The Railroad Retirement Board does not collect information on race and ethnicity. Similarly, there is no requirement to report information regarding the race and ethnicity of plan members in the Medicare Health Plan. There are also enrollees with an assigned race.

[24] The Social Security Administration treated Hispanic as a separate race category.

In past DA estimates production, the Medicare race categories were recoded to represent Black and non-Black. Records with a classification of Unknown on the Medicare file were prorated to Black and non-Black.[25] For DA in 2010, we used the Census Bureau's Person Characteristics File (PCF) and a model-based race to supplement the race information from the MEDB. Over the last decade, the Census Bureau has built a person-level file from administrative records primarily for application to decennial census research and development. The SSA Numident file (a file of Social Security numbers) is processed to produce a file of unique SSNs which are then replaced by randomly assigned Protected Identity Keys (PIKS) to maintain confidentiality (Farber and Miller, 2002). This file (the 100-percent Census Numident file) is then enhanced with demographic data to create the PCF (see Miller, Judson, and Sater, 2000 for a more complete description).

The MEDB-universe was matched to the PCF. This allowed us to use race information from three different sources: Medicare, Census 2000, and a model-based approach (see Resnick, 2002 for a more detailed discussion of the use of administrative records to predict race and Hispanic origin).

When different options for using these three sources were examined, the outcomes were very similar. Each option used the race from each source in a different order. The option used for DA in 2010 gives precedence to the race maintained in the MEDB and is outlined in the steps below:

1. If the Medicare record is White, Black, Asian, American Indian, or Other, the Medicare race is used.[26]
2. If the Medicare record is Hispanic, Unknown, or Blank, the Census race is used if it is White, Black, Asian, Pacific Islander, or American Indian.
3. If the Census race in step 2 is Multiple race, Unknown or Blank, the modeled race is used.

[25] The Black population was obtained as Black + (Black/(Black + non-Black)*Unknown). Similarly, non-Black = non-Black + (non-Black/Black + non-Black)*Unknown.

[26] Medicare race is collected from the Social Security card application. When most of the Medicare enrollees in our universe applied for an SSN, the race categories were White, Black and Other, with Other representing primarily Asian. For DA in 2010, the Other category was combined with the non-Black population.

4. The records that end up being modeled as Multiple race or Blank as a result of step 3 are distributed proportionally among the race categories (White, Black, Asian, American Indian or Other.)
5. The White, Asian, American Indian, and Other categories are considered non-Black.

The factors for those not enrolled are developed for the Black and non-Black race categories using the race as reported in the CPS. Multiracial respondents in the CPS were considered non-Black. These adjustment factors are then applied to the race distribution obtained using the steps just described.

ESTIMATES BY HISPANIC ORIGIN

As described earlier, in previous decades the production of DA estimates was limited to two broad race categories, Black and non-Black, with no reference to ethnicity. As the racial and ethnic composition of the nation's population has changed, there has been an increased interest in expanding beyond the DA Black and non-Black race categories. The Hispanic population has grown substantially over the past several decades. In 2009, Hispanics comprised an estimated 15.8 percent of the population (U.S. Census Bureau, 2010), up from 12.5 percent in 2000 (U.S. Census Bureau, 2001b), 9.0 percent in 1990 (U.S. Census Bureau 1991), and 6.4 percent in 1980 (Gibson and Jung, 2005). The growth in the Hispanic population has played an increasingly important role in shaping the size and age structure of the U.S. population. For 2010 DA, we have developed estimates of the population under 20 years of age by Hispanic origin, age, and sex for comparison with 2010 Census counts. The DA estimates are limited to these ages because of the available historical components of change (births, deaths, and estimates of international migration).

Historical availability of vital statistics data by Hispanic origin

NCHS began working with states to report vital statistics birth data by parental Hispanic origin in the 1970s. By 1988, 30 states and the District of Columbia reported births by Hispanic origin. In 1989, the U.S. Standard Birth Certificate was revised to formally incorporate questions on parental Hispanic origin. In 1990, the first year that a birth would be included in the DA estimates by

Hispanic origin; all but two states reported this information. Oklahoma added the Hispanic origin question to its birth certificate in 1991, and New Hampshire in 1993.

Death certificates also started to include a question on Hispanic origin in the 1970s. By 1980, 21 states had a Hispanic origin item on their death certificates. However, Florida (a state with a large Hispanic population) did not include this item until 1989, and California had more than 50 percent missing responses until 1983 (Arias et al., 2008). The Hispanic origin item became part of the U.S. Standard Death Certificate in the 1989 revision. In 1990, all states except Louisiana (1991), New Hampshire (1993), and Oklahoma (1997) reported deaths by Hispanic origin of the decedent.

Births by Hispanic origin

The Hispanic origin question on birth certificates provides information on mother's Hispanic origin and father's Hispanic origin. As with race, information is not recorded on the Hispanic origin of the child. To classify each birth as Hispanic or non-Hispanic, we first impute parental Hispanic origin for records where mother's and/or father's Hispanic origin is missing. We then classify children into one of the two categories, Hispanic or non-Hispanic, using information on reported parental Hispanic origin.

Imputing missing Hispanic origin of mother and/or father

Missing Hispanic origin of mother and/or father is proportionally imputed using state-level information from records where parental Hispanic origin is known. Two donor datasets are created. The first consists of vital statistics records from 1990-2000, and the second consists of records from 2000-2007. Both donor datasets are restricted to birth records where both mother and father's Hispanic origin are known. The 1990-2000 dataset serves as the donor for imputing Hispanic origin for births from April 1990 through March 2000, and the 2000-2007 dataset serves as the donor for imputing Hispanic origin from April 2000 through December 2007. Births for January 2008 through March 2010 are projections and therefore are not part of the imputation process. Missing Hispanic origin is imputed by state using proportions from the donor dataset. For example, in cases where mother is Hispanic and the Hispanic origin of father is unknown, fathers are imputed as Hispanic using the state-level proportion of fathers in the donor dataset who are Hispanic when mother is Hispanic. If the Hispanic origin of both mother and father is unknown, mother's Hispanic origin is imputed using the state-level proportion of all mothers who are Hispanic, and then father's Hispanic origin is imputed through proportional imputation using the imputed mother's Hispanic origin classification.

In the data used to construct estimates by Hispanic origin (1990 to 2007), approximately 787 thousand births (1.1 percent) were missing Hispanic origin of mother and approximately 11 million births (15.2 percent) were missing Hispanic origin of father. Among births missing Hispanic origin of mother or father, 669 thousand (6.1 percent) were missing both Hispanic origin of mother and father. Overall, 16.1 percent of mothers with an unknown Hispanic origin were imputed as Hispanic, and 18.0 percent of fathers.

Classifying births as Hispanic or non-Hispanic

The next step is to classify each birth as Hispanic or non-Hispanic using information on the Hispanic origin of mother and father. We create a low, middle, and high estimate of Hispanic births using three separate classification rules. The total number of births is consistent under each assumption; the only difference is the distribution of births by Hispanic origin. Under the high assumption, children are classified as Hispanic if mother or father is reported as Hispanic on the birth certificate. This provides the maximum number of possible Hispanic births. Under the low

assumption, Hispanic births are estimated as births where both parents are Hispanic plus 50 percent of births to one Hispanic and one non-Hispanic parent.

For the middle estimate, we use Census 2000 data on reported Hispanic origin of parents and children to proportionally classify children as Hispanic or non-Hispanic (similar to how race is assigned). Based on Census Bureau tabulations that included imputations when parental Hispanic origin was missing and projections from January 2008 through March 2010, 24.8 percent of births where at least one parent was Hispanic consisted of one Hispanic parent and one non-Hispanic parent. Previous research has found that the proportion of children with both a Hispanic and non-Hispanic parent who are reported as Hispanic is around 63 percent (Lee and Edmonston, 2005). The Census 2000 Kid Link file discussed earlier was also used to develop the proportions used in DA. The results were consistent with the findings from the previous research. As with the proportions developed for the assignment of Black and non-Black, cases where Hispanic origin of either parent or of child was edited were excluded. Overall, we found that 61.4 percent of children were reported as Hispanic when the mother was Hispanic and the father was non-Hispanic, and 69.8 percent of children were reported as Hispanic when the mother was non-Hispanic and the father was Hispanic. These proportions were applied to impute Hispanic origin of children in the birth data when one-parent is Hispanic and the other is non-Hispanic. Overall, 65.6 percent of births with one-Hispanic parent were classified as Hispanic in the middle estimate. When both parents are of the same ethnicity (Hispanic or non-Hispanic) all of the births were assigned the ethnicity of their parents.

The DA estimate for Hispanics aged 0 to 19 includes 81.3 million births that occurred on or after April 1, 1990. About 20.4 percent of births are classified as Hispanic under the low assumption, 21.3 percent under the middle assumption, and 23.3 percent under the high assumption. The proportion of births that are classified as Hispanic is higher for the younger ages than among the older ages, which reflects an increase in births to Hispanic parents over time. For example, in the middle estimate 25.8 percent of births in the cohort aged 0 to 4 as of April 1, 2010 are classified as Hispanic, compared to 16.8 percent of births in the 15 to 19 year old cohort.

Deaths by Hispanic origin

Deaths are processed for the Hispanic origin series using similar methods and imputation procedures as those used to produce estimates by Black and non-Black. There are just over 12 thousand deaths missing information on Hispanic origin (1.82 percent of deaths). To impute Hispanic origin for deaths with missing values, the dataset is sorted by birth year, birth month, state of residence, race, and sex. Hispanic origin is then imputed using the reported Hispanic origin from a record with similar characteristics as indicated by the sort sequence. Overall, 10.9 percent of deaths with a missing Hispanic origin are imputed as Hispanic.

Only one series of estimated deaths by Hispanic origin was used. The estimate of deaths for the cohort that would be age 0 to 19 as of April 1, 2010 is approximately 780 thousand. Of these deaths, 14.0 percent are Hispanic.

Projecting births and deaths through March 31, 2010

The same method that is used to produce estimates of births and deaths by Black/non-Black (see page 12) for the period from January 2008 through March 2010 is used to estimate the births and deaths by Hispanic/non-Hispanic for this period.

International migration

Net international migration from April 1990 through March 2000 is estimated using the methods from 2000 DA. Although the Census Bureau did not produce 2000 Demographic Analysis Population Estimates (DAPE) by Hispanic origin, the international migration components were estimated by Hispanic origin.[27] For 2000 to 2010, net international migration components were developed by race and Hispanic origin using the methods described earlier.

[27] For more information on the methods used to estimate international migration from April 1990 through March 2000, please see the Population Division Working Papers (58 through 64) online at <http://www.census.gov/population/www/techpap.html>.

OUTREACH TO EXTERNAL EXPERTS

The Census Bureau has a long history of collaborating with external experts on DA. Two sets of recommendations were considered in the development of the overall strategy for DA in 2010. The first set of recommendations came from the Census 2000 Monitoring Board (U.S. Census 2000 Monitoring Board, 2001). The second set of recommendations came from a 2008 report from the National Research Council of the National Academies, "Coverage Measurement in the 2010 Census" (National Research Council, 2008).

While most of the recommendations from the Census 2000 Monitoring Board pertain to how the DA estimates of coverage are to be used, they also include a call for the Census Bureau to increase its capacity to measure immigration and an evaluation of the DA program prepared by Passel (2001: 86).

Recommendations for DA from the National Research Council of the National Academies were provided by a panel on "Coverage Evaluation and Correlation Bias in the 2010 Census." The recommendations included working to improve the measurement of international migration and identifying ways to assess the uncertainty in the DA estimates (National Research Council, 2008).

In addition to carefully considering prior recommendations, we also sought to obtain input from external experts throughout the 2010 DA effort. This has involved presenting the plans for DA in 2010 to a wide range of audiences and inviting key experts from both inside and outside the federal statistical community to participate in a DA workshop on January 8, 2010. Several participants from this workshop have continued their collaboration with us through participating in panel sessions at professional conferences and by working with us at the Census Bureau as part of the Census Bureau's "Summer at Census" program.

SUMMARY

The goal of the 2010 DA program has been to provide the best national estimates possible by age, sex, and the DA race and ethnicity categories to serve as the foundation for the demographic analysis of the 2010 Census counts. We also considered it essential to communicate the uncertainty in these estimates due to the limitations of the available data. This paper has provided a description of the development of alternative estimates for each component of the DA methodology that were used in the five series of DA estimates. These five series of estimates were constructed to help communicate how the uncertainty in each of the components of DA translates into uncertainty in the estimates by age, sex, and in the DA race and ethnicity categories. The alternative estimates for each component have been provided in this paper prior to the dissemination of the DA estimates to allow for a more informed interpretation of the range of DA estimates when they are released on December 6, 2010. We will continue to work on improving the DA estimates and will incorporate the final vital statistics and Medicare data. Revised estimates will be disseminated as they become available.

REFERENCES

Adlakha, Arjun L., J. Gregory Robinson, and Amy Symens Smith. 2002. "Alternative Rules for Assigning Race of Birth: Effect on Birth Totals, Implications for Vital Rates and Census Undercount Estimates by Race." Paper presented at the Annual Meeting of the Southern Demographic Association. Austin, Texas, October 10-12.

Arias, Elizabeth, William S. Schauman, Karl Eschbach, Paul D. Sorlie, and Eric Backlund. 2008. "The Validity of Race and Hispanic Origin Reporting on Death Certificates in the United States." National Center for Health Statistics. Vital Health Stat 2(148).

Coale, Ansley J. 1955. "The Population of the United States in 1950 Classified by Age, Sex and Color--A Revision of Census Figures." *Journal of the American Statistical Association* 50(269):16-54.

Das Gupta, Prithwis. 1991. "Models for Assessing Errors in Undercount Rates Based on Demographic Analysis." DA Evaluation Project D11. Preliminary Research and Evaluation Memorandum No. 84. U.S. Census Bureau, Washington, D.C.

Farber, J. and E. Miller. 2002. "Matching Census 2000 to Administrative Records." Proceedings of the 2002 Joint Statistical Meeting, Alexandria, Va.

Gibbs, Jim C., Gregory S. Harper, Marc J. Rubin, and Hyon B. Shin. 2001. "Evaluating Components of International Migration: Native Emigrants." Population Division Working Paper Series No. 63. U.S. Census Bureau, Washington, D.C.

Gibson, Campbell and Kay Jung. 2005. "Historical Census Statistics on Population Totals by Race 1790 to 1990, and by Hispanic Origin, 1970 to 1990, for Large Cities and Other Urban Places in the United States." Population Division Working Paper Series No. 76. U.S. Census Bureau, Washington, D.C.

Himes, C. L. And C.C. Clogg. 1992. "An Overview of Demographic Analysis as a Method for Evaluating Census Coverage in the United States." *Population Index* 58(4): 587-607.

Hogan, Howard, William Bell, Lynn Weidman, and Allen Schirm, 2003. "Integrating Survey, Demographic, and Modeling Methods." Joint Statistical Meetings, Section on Government Statistics.

Jones, Nicholas A. and Amy Symens Smith. 2003. "New Explorations of Race Reporting for Interracial Couples and Their Children: Census 2000." Paper presented at the Annual Meeting of the Population Association of America. Minneapolis, Minnesota, May 2, 2003.

Lee, Sharon M. and Barry Edmonston. 2005. "New Marriages, New Families: U.S. Racial and Hispanic Intermarriage." Population Bulletin, 60(2): Washington D.C.

MacDorman, Marian, Fay Menacker, and Eugene Declercq. 2010. "Trends and Characteristics of Home and Other Out-of-Hospital Births in the United States, 1990-2006." National Vital Statistics Report; 58(11). National Center for Health Statistics.

McDevitt, Tom, Martin O'Connell, and Colleen Joyce. 2001. "Evaluating the Components of Births and Deaths used in Demographic Analysis." Population Division, U.S. Census Bureau, Washington, D.C.

Miller, E.R., D.H. Judson, and D. Sater. 2000. "The 100% Census Numident: Demographic Analysis of Modeled Race and Hispanic Origin Estimates Based Exclusively on Administrative Records Data." Paper presented at the Annual Meeting of the Southern Demographic Association, New Orleans, Louisiana, October 26-28.

National Research Council. 2008. *Coverage Measurement in the 2010 Census*. Panel on Coverage Evaluation and Correlation Bias in the 2010 Census, Robert M. Bell and Michael L. Cohen

(Eds.), Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
____. 1994. *Counting People in the Information Age*. Panel to Evaluate Alternative Census Methods, Duane L. Steffey and Norman M. Bradburn, eds., Committee on National Statistics. Washington, D.C.: National Academy Press.
____. 1985. The *Bicentennial Census: New Directions for Methodology in 1990*. Panel on Decennial Census Methodology, Constance F. Citro and Michael L. Cohen, eds., Committee on National Statistics. Washington, D.C.: National Academy Press.
Passel, Jeffrey S. 2001. "Demographic Analysis: An Evaluation." Section 4 in U.S. Census Monitoring Board, Final Report to Congress, 2001.
____. 1992. "Alternative Estimates of Black Births Corrected for Underregistration: 1935-1980." The Urban Institute, Washington, D.C.
Passel, Jeffrey S. and J. Gregory Robinson. 1988. Methodology for Developing Estimates of Coverage in the 1980 Census Based on Demographic Analysis: Population Aged 65 and older in 1980. Preliminary Evaluation Results Memorandum No 116.
Preston, Samuel H., Irma T. Elo, Andrew Foster, and Haishan Fu, 1998. "Reconstructing the Size of the African American Population by Age and Sex, 1930–1990." *Demography* 35(1): 1–21.
Resnick, D.M. 2002. "Improving the Accuracy of Race and Hispanic Origin in Administrative Records." Proceedings of the 2002 Joint Statistical Meeting, Alexandria, Va.
Robinson, J. Gregory. 2010. "Coverage of Population in Census 2000 Based on Demographic Analysis: The History Behind the Numbers." Paper prepared for the Demographic Analysis Technical Review Workshop, January.
____. 1991. "Demographic Analysis Evaluation Project D1: Error in the Birth Registration Completeness Estimates." Unpublished Census Bureau Report.
Robinson, J. Gregory, B. Ahmed, P. Das Gupta, and K. A. Woodrow. 1993. "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis." *Journal of the American Statistical Association* 88(423): 1061-1071.
Schachter, J. 2008. "Estimating Native Emigration from the United States," Memorandum dated December 24, developed during contract work for the U.S. Census Bureau.
Siegel, Jacob S. and Melvin Zelnik. 1966. "An Evaluation of Coverage in the 1960 Census of Population by Techniques of Demographic Analysis and by Composite Methods." Proceedings of the Social Statistics Section of the American Statistical Association.
Turner, Joanna, Michel Boudreaux and Victoria Lynch. 2009. "A Preliminary Evaluation of Health Insurance Coverage in the 2008 American Community Survey." U.S. Census Bureau, September.
U.S. Census Bureau. 2010. Table 3. Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States: April 1, 2000 to July 1, 2009 (NC-EST2009-03)
____. 2009. "Source and Accuracy Statement for the October 2008 CPS Microdata File on School Enrollment.", Memorandum, August.
____. 2001. "ESCAP II: Demographic Analysis Results," by J. Gregory Robinson. Executive Steering Committee on Accuracy and Coverage Evaluation Policy II, October.
____.2001b. Census of Population and Housing, 2000: Summary File 1. Table P8: Hispanic or Latino by Race.
____.1991. Census of Population and Housing, 1990: Summary Tape File 1. Table P009: Hispanic Origin.

____. 1988. “The Coverage of Population in the 1980 Census,” by Robert Fay, Jeffrey S. Passel, and J. Gregory Robinson. Evaluation and Research Reports, PHC80-E4.
____. 1974. “Estimates of Coverage of Population by Sex, Race, and Age: Demographic Analysis,” by Jacob S. Siegel. Evaluation and Research Reports, PHC(E)-4.
____. 1970. “Test of Birth Registration Completeness, 1964 to 1968.” Evaluation and Research Program, PHC(E)-2.
U.S. Census 2000 Monitoring Board. 2001. Final Report to Congress. www.cmbp.gov
U.S. Department of Health and Human Services, National Center for Health Statistics. 2008. *Detailed Technical Notes, 2006*.
U. S. House of Representatives Committee on Ways and Means. 1998. *1998 Green Book - Background Material and Data on Programs within the Jurisdiction of the Committee on Ways and Means.* Washington, DC: U.S. Government Printing Office.

APPENDIX A. HISTORICAL DEVELOPMENT OF THE DA METHODOLOGY

Demographic Analysis is an analytic approach that has been used to develop estimates for comparison with the national population in every census since 1950 (see Coale, 1955: Siegel and Zelnik, 1966; U.S. Census Bureau, 1974, 1988; and Robinson, 2010 for the demographic evaluations of the 1950-2000 censuses). The current DA methodology has developed from the use of a collection of analytical and estimation techniques focused on different subgroups of the population. As various data sources became available, or were considered of reliable quality, the analytical techniques were replaced with comparisons with direct estimates of the population. The current DA methodology allows for the direct estimation of the national population for all ages by sex and race (Black and non-Black).

The DA methodology has continued to evolve as new data sources and later years of vital statistics data have become available. Changes in the census questionnaire and the demographics of the nation have also played a role in the evolution of the DA methodology. The inclusion of the option to mark multiple races in Census 2000 and the 2010 Census questionnaires and increases in multiracial births pose methodological challenges for DA. While the data used for DA are essentially independent from the census, each census does bring new opportunities to re-assess the DA estimates, such as with Census 2000. In 2000, the initial DA estimate of 279.6 million was revised upward to 281.8 million, reflecting in large part a change in the assumptions about international migration during the 1990s. This change was based in part on the count of the foreign-born population in the census (U.S. Census Bureau, 2001).

External reviews of the DA methodology have been conducted each decade and have helped guide the evolution of the DA methodology. The National Research Council (1985, 1994, 2008) has issued reviews and recommendations the last two decades regarding the DA program. Passel (1992) reviewed the DA program at the time of the 1990 Census. In 1990, explicit measures of uncertainty in the DA net undercount estimates were developed for the first time (Das Gupta, 1991; Robinson et al., 1993). Himes and Clogg (1992) also provided a statistically-based assessment of the DA estimates. The U.S. Census 2000 Monitoring Board (2001) provided a review of the 2000 DA estimates, with a recommendation to increase the capacity to measure immigration.

Table 1. Difference of Demographic Analysis Estimate and Census Count of the Resident Population: April 1, 1940 to April 1, 2000

Population by race	(Numbers in thousands)						
	1940[1]	1950[1]	1960	1970	1980	1990	2000
Total							
DA estimate	139,678	157,863	185,024	208,955	229,347	252,876	281,760
Census count	132,165	151,326	179,323	203,302	226,546	248,710	281,422
Difference	-7,513	-6,537	-5,700	-5,653	-2,802	-4,166	-338
Percent difference	-5.38	-4.14	-3.08	-2.71	-1.22	-1.65	-0.12
Black							
DA estimate	14,052	16,271	20,199	24,154	27,940	32,265	37,443
Census count	12,866	15,046	18,872	22,581	26,683	30,483	36,404
Difference	-1,187	-1,225	-1,327	-1,573	-1,257	-1,782	-1,039
Percent difference	-8.44	-7.53	-6.57	-6.51	-4.50	-5.52	-2.78
Non-Black							
DA estimate	125,626	141,592	164,825	184,801	201,407	220,610	244,317
Census count	119,300	136,280	160,451	180,721	199,862	218,227	245,018
Difference	-6,326	-5,312	-4,374	-4,079	-1,545	-2,384	701
Percent difference	-5.04	-3.75	-2.65	-2.21	-0.77	-1.08	0.29

[1] Census counts for 1940 and 1950 include Alaska and Hawaii.

Sources: 1990 and 2000 DA estimates are published in *Census 2000 ESCAP II report on DA* , 1990-Appendix table B4; 2000-Appendix Table B3: http://www.census.gov/dmd/www/pdf/Report1.PDF. 1940 to 1980 DA estimates are consistent with estimates published in: Robinson, J. Gregory, B. Ahmed, P. Das Gupta, and K. A. Woodrow. 1993a. "Estimation of Population Coverage in the 1990 United States Census Based on Demographic Analysis." Journal of the American Statistical Association 88(423): 1061-1071.

Table 2. 2010 Demographic Analysis Estimates Matrix				
	Births	**Deaths**	**International Migration**	**Medicare Enrollment**
DA Estimate	**Cohorts born between 1945 and 2010**			**Cohorts born before 1945**
Low Series	**(A) Low Births Series**: Reduced correction factors for underregistration through 1984. 1980 to 2010 race based on Kid Link-Black alone proportions. Race for births between 1945 and 1980 are based on race of father.	**(A) Registered Deaths**: Infant deaths corrected for underregistration through 1959.	**(A-1) Low NIM Series**: *Foreign-Born Immigration* - Lower bound Residence One Year Ago *Foreign-Born Emigration* - High estimate using residual method *Net Native-Born Migration* - Schachter methodology *Net Migration from Puerto Rico* - Residence One Year Ago Method *Residual Foreign Born (2000)* - Low Coverage assumption	**(A) Low Medicare Series**: Lower bound confidence interval (90%) using delayed enrollment factors from the Medicare Enrollment file and never-enrolled factors from the CPS (2002-2008).
Low Middle Series	**(B) Middle Births Series**: Corrections for underregistration through 1984. 1980 to 2010 race based on Kid Link-Black alone proportions. Race for births between 1945 and 1980 are based on race of father.		**(A-2) Low NIM Series**: *Foreign-Born Immigration* - Lower bound Residence One Year Ago *Foreign-Born Emigration* - High estimate using residual method *Net Native-Born Migration* - Schachter methodology *Net Migration from Puerto Rico* - Residence One Year Ago Method *Residual Foreign Born (2000)* - 85% Coverage assumption	**(B) Middle Medicare Series**: Estimates developed using delayed enrollment factors from the Medicare Enrollment file and never-enrolled factors from the CPS (2002-2008).
Middle Series			**(B) Middle NIM Series**: *Foreign-Born Immigration* - Residence One Year Ago *Foreign-Born Emigration* - Residual method *Net Native-Born Migration* - Schachter methodology *Net Migration from Puerto Rico* - Residence One Year Ago Method *Residual Foreign Born (2000)* - 85% Coverage assumption *Born Abroad of American Citizen Parents* - Middle estimate	
High Middle Series			**(C) High Middle NIM Series**: *Foreign-Born Immigration* - Year of Entry *Foreign-Born Emigration* - Residual method *Net Native-Born Migration* - Schachter methodology *Net Migration from Puerto Rico* - Residence One Year Ago Method *Residual Foreign Born (2000)* - 85% Coverage assumption *Born Abroad of American Citizen Parents* - Middle estimate	
High Series	**(C) High Births Series**: Increased correction factors for underregistration through 1984. 1980 to 2010 race based on Kid Link-Black alone proportions. Race for births between 1945 and 1980 are based on race of father.		**(D) High NIM Series**: *Foreign-Born Immigration* - Year of Entry with coverage factors *Foreign-Born Emigration* - Low estimate using residual method *Net Native-Born Migration* - Schachter methodology *Net Migration from Puerto Rico* - Residence One Year Ago Method *Residual Foreign Born (2000)* - 85% Coverage assumption *Born Abroad of American Citizen Parents* - High estimate	**(C) High Medicare Series**: Upper bound confidence interval (90%) using enrollment factors from the CPS (2002-2008).

Table 3. Demographic Analysis and NCHS Birth Data for the United States, 1935-2010

Age on April 1, 2010	Census Year of Birth	DA 2010 Birth Series			Registered Births DA 2010	Registered Births NCHS[1]
		Low	Middle	High		
	Total 1935-2010	275,363,282	276,355,807	277,348,330	273,047,394	260,372,171[2]
	Total 1945-2010	249,419,990	249,890,731	250,361,470	248,321,595	235,663,471[2]
0	2009-10	4,135,000	4,135,000	4,135,000	4,135,000	N/A[3]
1	2008-09	4,210,089	4,210,089	4,210,089	4,210,089	N/A[3]
2	2007-08	4,318,577	4,318,577	4,318,577	4,318,577	N/A[3]
3	2006-07	4,291,898	4,291,898	4,291,898	4,291,898	4,291,898
4	2005-06	4,164,249	4,164,249	4,164,249	4,164,249	4,164,249
5	2004-05	4,107,394	4,107,394	4,107,394	4,107,394	4,107,394
6	2003-04	4,111,055	4,111,055	4,111,055	4,111,055	4,111,055
7	2002-03	4,029,542	4,029,542	4,029,542	4,029,542	4,029,542
8	2001-02	4,014,671	4,014,671	4,014,671	4,014,671	4,014,671
9	2000-01	4,048,193	4,048,193	4,048,193	4,048,193	4,048,193
10	1999-00	3,997,766	3,997,766	3,997,766	3,997,766	3,997,766
11	1998-99	3,943,755	3,943,755	3,943,755	3,943,755	3,943,755
12	1997-98	3,898,417	3,898,417	3,898,417	3,898,417	3,898,417
13	1996-97	3,882,831	3,882,831	3,882,831	3,882,831	3,882,831
14	1995-96	3,898,606	3,898,606	3,898,606	3,898,606	3,898,606
15	1994-95	3,930,609	3,930,609	3,930,609	3,930,609	3,930,609
16	1993-94	3,992,092	3,992,092	3,992,092	3,992,092	3,992,092
17	1992-93	4,045,919	4,045,919	4,045,919	4,045,919	4,045,919
18	1991-92	4,111,537	4,111,537	4,111,537	4,111,537	4,111,537
19	1990-91	4,148,094	4,148,094	4,148,094	4,148,094	4,148,094
20	1989-90	4,078,732	4,078,732	4,078,732	4,078,732	4,078,732
21	1988-89	3,939,637	3,939,637	3,939,637	3,939,637	3,939,637
22	1987-88	3,833,983	3,833,983	3,833,983	3,833,983	3,833,983
23	1986-87	3,761,451	3,761,451	3,761,451	3,761,451	3,761,451
24	1985-86[4]	3,765,652	3,765,652	3,765,652	3,765,652	3,765,652
25	1984-85	3,688,406	3,688,831	3,689,254	3,687,420	3,687,420
26	1983-84	3,632,959	3,633,796	3,634,630	3,631,008	3,631,008
27	1982-83	3,688,621	3,689,887	3,691,156	3,685,663	3,685,663
28	1981-82	3,644,100	3,645,766	3,647,431	3,640,214	3,640,214
29	1980-81	3,617,338	3,619,395	3,621,454	3,612,536	3,612,536
30	1979-80	3,535,137	3,537,447	3,539,756	3,529,749	3,529,596
31	1978-79	3,376,056	3,378,569	3,381,083	3,370,190	3,370,183
32	1977-78	3,324,881	3,327,844	3,330,805	3,317,972	3,317,980
33	1976-77	3,223,132	3,226,434	3,229,738	3,215,422	3,215,422
34	1975-76	3,151,715	3,155,250	3,158,785	3,143,466	3,143,459
35	1974-75	3,175,855	3,179,681	3,183,507	3,166,929	3,166,935
36	1973-74	3,122,766	3,126,845	3,130,922	3,113,252	3,113,253
37	1972-73	3,233,513	3,238,100	3,242,687	3,222,808	3,222,809
38	1971-72	3,476,509	3,481,912	3,487,317	3,463,900	3,463,889
39	1970-71	3,758,758	3,764,783	3,770,807	3,744,700	3,744,700
40	1969-70	3,643,575	3,649,838	3,656,102	3,628,961	3,628,968
41	1968-69	3,552,993	3,559,757	3,566,520	3,537,209	3,537,202
42	1967-68	3,504,884	3,512,132	3,519,383	3,487,969	3,487,962
43	1966-67	3,612,442	3,620,469	3,628,493	3,593,715	3,593,737
44	1965-66	3,735,965	3,745,707	3,755,448	3,713,236	3,713,230
45	1964-65	3,987,314	3,997,894	4,008,477	3,962,625	3,962,638
46	1963-64	4,116,374	4,127,405	4,138,435	4,090,637	4,090,665
47	1962-63	4,171,886	4,183,515	4,195,143	4,144,752	4,145,003
48	1961-62	4,269,408	4,281,846	4,294,283	4,240,389	4,240,082
49	1960-61	4,304,833	4,318,043	4,331,253	4,274,008	4,274,024
50	1959-60	4,284,717	4,298,559	4,312,401	4,252,417	4,256,738
51	1958-59	4,278,709	4,293,296	4,307,884	4,244,672	4,245,979
52	1957-58	4,314,818	4,330,259	4,345,701	4,278,789	4,279,015
53	1956-57	4,228,719	4,244,816	4,260,911	4,191,161	4,190,905
54	1955-56	4,152,235	4,169,201	4,186,170	4,112,645	4,112,742
55	1954-55	4,093,165	4,111,378	4,129,589	4,050,670	4,050,618
56	1953-54	3,993,169	4,012,228	4,031,289	3,948,695	3,948,723
57	1952-53	3,920,437	3,940,565	3,960,693	3,873,473	3,873,875
58	1951-52	3,851,979	3,873,865	3,895,748	3,800,917	3,800,855
59	1950-51	3,661,300	3,684,646	3,707,996	3,606,818	3,606,497
60	1949-50	3,628,625	3,654,383	3,680,141	3,568,523	3,568,535
61	1948-49	3,609,812	3,638,238	3,666,661	3,543,489	3,543,805
62	1947-48	3,704,510	3,736,154	3,767,796	3,630,675	3,630,665
63	1946-47	3,696,119	3,730,603	3,765,089	3,615,656	3,615,288
64	1945-46	2,792,507	2,821,645	2,850,783	2,724,516	2,724,570
65	1944-45	2,882,475	2,915,716	2,948,958	2,804,912	2,803,527
66	1943-44	2,968,195	3,005,526	3,042,856	2,881,091	2,879,326
67	1942-43	3,015,495	3,057,883	3,100,272	2,916,590	2,914,447
68	1941-42	2,682,503	2,727,555	2,772,607	2,577,379	2,575,351
69	1940-41	2,516,089	2,566,306	2,616,522	2,398,919	2,396,983
70	1939-40	2,423,559	2,480,000	2,536,441	2,291,863	2,290,188
71	1938-39	2,436,517	2,498,300	2,560,082	2,292,356	2,290,680
72	1937-38	2,391,951	2,456,162	2,520,373	2,242,128	2,240,513
73	1936-37	2,299,377	2,363,375	2,427,374	2,150,044	2,148,579
74	1935-36	2,327,131	2,394,253	2,461,375	2,170,517	2,169,106

[1] NCHS registered births for Alaska are included beginning in 1945 when data became available

[2] Total for ages 3 to 74

[3] Data are not available

[4] Registration is considered 100% complete from this year forth

Note: Low and High series come from a 30% decrease or increase in the underregistered births

Sources:

Vital Statistics of the United States, 1935-1996

National Vital Statistics Reports, 1997-2007

U.S. Census Bureau, Demographic Analysis Tabulations for Birth Data

Table 4. Census 2000 Residual Percent Coverage Profile by Hispanic Origin, Age, and Sex

Hispanic origin and age	Male	Female
Hispanic		
0-15	92.5	92.5
16-29	90.0	90.0
30-49	85.0	92.5
50+	95.0	95.0
Non-Hispanic		
0-15	95.0	95.0
16-29	92.5	92.5
30-49	90.0	95.0
50+	100.0	100.0

Note: This coverage profile was developed by the U.S. Census Bureau from ACE Revision II and CPS data.
Source: U.S. Census Bureau

Table 5. Net International Migration, All Ages: 2000 to 2010

	Foreign-Born Immigration					Foreign-Born Emigration			Net Native Migration	Born Abroad of U.S. Citizen Parents (Net)		PR - US Migration
Period	ROYA[1]	YOE[2]	ROYA: Lower Bound	YOE: Upper Bound	YOE with Representation Factors Applied	Total	Total: High	Total: Low	Foreign Census Method	Change in Stock	Change in Stock: No Citizen Countries	ROYA: Net
4/1/00 - 3/31/10	11,853,066	13,475,814	11,418,920	13,927,832	14,236,273	2,281,258	2,567,081	1,982,369	-452,280	624,240	397,052	204,871
4/1/00 - 6/30/00	357,731	402,642	341,873	420,352	431,623	52,146	58,970	45,100	-11,307	15,606	9,926	2,783
7/1/00 - 6/30/01	1,430,922	1,610,568	1,367,492	1,681,410	1,726,490	208,584	235,880	180,398	-45,228	62,424	39,705	11,385
7/1/01- 6/30/02	1,367,448	1,555,521	1,305,137	1,624,927	1,660,958	209,254	236,690	180,875	-45,228	62,424	39,705	11,637
7/1/02- 6/30/03	1,234,564	1,450,864	1,178,284	1,513,873	1,541,937	213,652	241,940	184,388	-45,228	62,424	39,705	11,889
7/1/03 - 6/30/04	1,135,595	1,308,038	1,082,769	1,367,382	1,381,272	219,244	248,404	189,033	-45,228	62,424	39,705	12,141
7/1/04- 6/30/05	1,196,231	1,332,728	1,163,339	1,365,606	1,448,445	226,646	257,048	195,082	-45,228	62,424	39,705	12,393
7/1/05 - 6/30/06	1,197,655	1,383,708	1,168,713	1,411,800	1,433,505	236,861	267,205	205,048	-45,228	62,424	39,705	33,685
7/1/06- 6/30/07	1,122,326	1,309,276	1,087,248	1,343,178	1,378,169	243,540	274,493	210,983	-45,228	62,424	39,705	27,579
7/1/07 - 6/30/08	1,077,000	1,196,610	1,046,576	1,226,644	1,250,778	243,273	274,495	210,399	-45,228	62,424	39,705	34,010
7/1/08 - 6/30/09	990,625	1,100,491	958,565	1,127,234	1,133,198	237,163	263,448	209,155	-45,228	62,424	39,705	27,068
7/1/09 - 3/31/10	742,969	825,368	718,924	845,426	849,899	190,895	208,508	171,908	-33,921	46,818	29,781	20,301

[1] ROYA = Residence One Year Ago

[2] YOE = Year of Entry

Source: U.S. Census Bureau

Table 6. Medicare-Based Population Estimates for the Population 65 and Over by Race and Sex

Series and Measure	Total	Black		Non-Black	
		Male	Female	Male	Female
Number of Medicare Enrollees[1]	37,450,874	1,243,634	1,925,162	14,876,114	19,405,964
Series 1[2]					
Upper Bound	40,101,276	1,409,808	2,111,738	16,002,784	20,576,946
Estimate	39,956,699	1,387,865	2,088,122	15,952,770	20,527,942
Lower Bound	39,813,849	1,366,595	2,065,030	15,903,059	20,479,165
Series 2[3]					
Upper Bound	40,876,432	1,438,783	2,151,468	16,379,473	20,906,708
Estimate	40,726,013	1,415,943	2,126,964	16,327,019	20,856,087
Lower Bound	40,577,431	1,393,816	2,103,008	16,274,901	20,805,706

[1] Number of Medicare Enrollees is for April 1, 2009 (before adjustments for underenrollment and projections to April 1, 2010)

[2] Series 1 = Medicare-Based Delayed Enrollment and CPS Never Enrolled Factors from the 2000s

[3] Series 2 = CPS Underenrollment Factors Only

Source: U.S. Census Bureau

Table 7. Component- and Series-Based Estimates of Implied Percent Underenrollment in Medicare by Race and Sex for Selected Ages: 2009

Race and age in 2009	Male			Female		
	DA Component	Series 1[1]	Series 2[2]	DA Component	Series 1[1]	Series 2[2]
Black						
65+	-	6.04	9.10	-	5.46	7.67
65-69	12.00	11.16	13.70	8.20	9.67	12.10
70-74[3]	11.07	5.43	7.60	5.87	5.06	6.80
75-79	-	4.24	5.20	-	3.95	5.00
80-84	-	3.80	5.20	-	3.55	5.00
85+	-	3.67	5.20	-	3.45	5.00
Non-Black						
65+	-	3.47	6.13	-	3.24	4.88
65-69	4.50	6.54	11.20	3.40	5.62	9.60
70-74[3]	4.04	3.19	4.40	2.69	3.10	4.00
75-79	-	2.39	2.90	-	2.39	2.50
80-84	-	2.05	2.90	-	2.10	2.50
85+	-	1.97	2.90	-	2.03	2.50

[1] Series 1 = Medicare-Based Delayed Enrollment and CPS Never Enrolled Factors from the 2000s

[2] Series 2 = CPS Underenrollment Factors Only

[3] Component-based estimates are for ages 70-73.

Source: U.S. Census Bureau

Table 8. Percent of Children in Selected Race Groups by Parental Race Combination

Race of parents			Race of child (percent of total children)			
Mother	Father	Total children	Black alone	Black in combination	Not Black alone or in combination	Total
Black	Black	3,581,015	99.53	0.17	0.30	100.00
Black	White	92,150	33.77	39.82	26.41	100.00
Black	AIAN	4,971	58.26	16.40	25.35	100.00
Black	Asian	5,128	34.91	40.87	24.22	100.00
Black	NHPI	994	43.86	34.91	21.23	100.00
Black	Asian and NHPI	6,294	36.19	40.32	23.48	100.00
Black	Black in Combination	17,927	62.07	36.36	1.57	100.00
White	Black	322,474	32.47	48.65	18.88	100.00
AIAN	Black	12,390	39.32	31.58	29.10	100.00
Asian	Black	27,017	31.59	56.79	11.62	100.00
NHPI	Black	3,814	33.67	51.34	15.00	100.00
Asian and NHPI	Black	31,640	31.56	56.46	11.98	100.00
Black in Combination	Black	34,994	53.73	44.33	1.95	100.00
White	White	34,641,749	0.05	0.05	99.89	100.00
AIAN	White	145,612	0.11	0.19	99.70	100.00
Asian	White	292,857	0.05	0.11	99.84	100.00
NHPI	White	15,231	0.22	0.24	99.55	100.00
Asian and NHPI	White	314,750	0.05	0.12	99.82	100.00
Black in Combination	White	24,414	2.27	57.14	40.60	100.00
White	AIAN	148,645	0.15	0.26	99.59	100.00
AIAN	AIAN	247,800	0.12	0.09	99.79	100.00
Asian	AIAN	3,061	0.75	0.26	98.99	100.00
NHPI	AIAN	502	0.80	0.40	98.80	100.00
Asian and NHPI	AIAN	3,702	0.73	0.27	99.00	100.00
Black in Combination	AIAN	916	8.95	50.11	40.94	100.00
White	Asian	138,472	0.09	0.18	99.73	100.00
AIAN	Asian	2,366	0.17	0.85	98.99	100.00
Asian	Asian	1,787,356	0.04	0.02	99.94	100.00
NHPI	Asian	3,026	0.07	0.23	99.70	100.00
Asian and NHPI	Asian	1,794,161	0.04	0.02	99.94	100.00
Black in Combination	Asian	1,685	2.55	52.11	45.34	100.00
White	NHPI	18,630	0.21	0.28	99.51	100.00
AIAN	NHPI	920	0.65	0.76	98.59	100.00
Asian	NHPI	4,048	0.02	0.10	99.88	100.00
NHPI	NHPI	53,804	0.14	0.08	99.77	100.00
Asian and NHPI	NHPI	59,329	0.13	0.09	99.78	100.00
Black in Combination	NHPI	264	8.33	62.12	29.55	100.00
White	Asian and NHPI	162,987	0.10	0.19	99.70	100.00
AIAN	Asian and NHPI	3,469	0.35	0.78	98.88	100.00
Asian	Asian and NHPI	1,796,080	0.04	0.02	99.94	100.00
NHPI	Asian and NHPI	57,970	0.14	0.10	99.76	100.00
Asian and NHPI	Asian and NHPI	1,869,652	0.04	0.02	99.94	100.00
Black in Combination	Asian and NHPI	2,052	3.17	54.34	42.50	100.00
White	Black in Combination	32,939	2.32	66.04	31.64	100.00
AIAN	Black in Combination	1,579	6.27	52.50	41.23	100.00
Asian	Black in Combination	4,105	2.53	53.69	43.78	100.00
NHPI	Black in Combination	464	6.68	66.81	26.51	100.00
Asian and NHPI	Black in Combination	4,804	2.83	56.27	40.90	100.00
Black in Combination	Black in Combination	32,995	7.04	88.08	4.88	100.00

AIAN = American Indian and Alaska Native Alone

NHPI = Native Hawaiian and Other Pacific Islander Alone

Notes: Race of parents refers to race alone. Populations for each race group include both Hispanics and non-Hispanics.

Source: U.S. Census Bureau